



# Lecture 6: Confidence

# Recap

# $\chi^2$ distribution

- $\chi^2$  distribution is the sum of  $N$  independent variables  $X_i$ 
  - Where the distributions  $X_i$  are distributed as normal
  - $N$  denotes the number of degrees of freedom

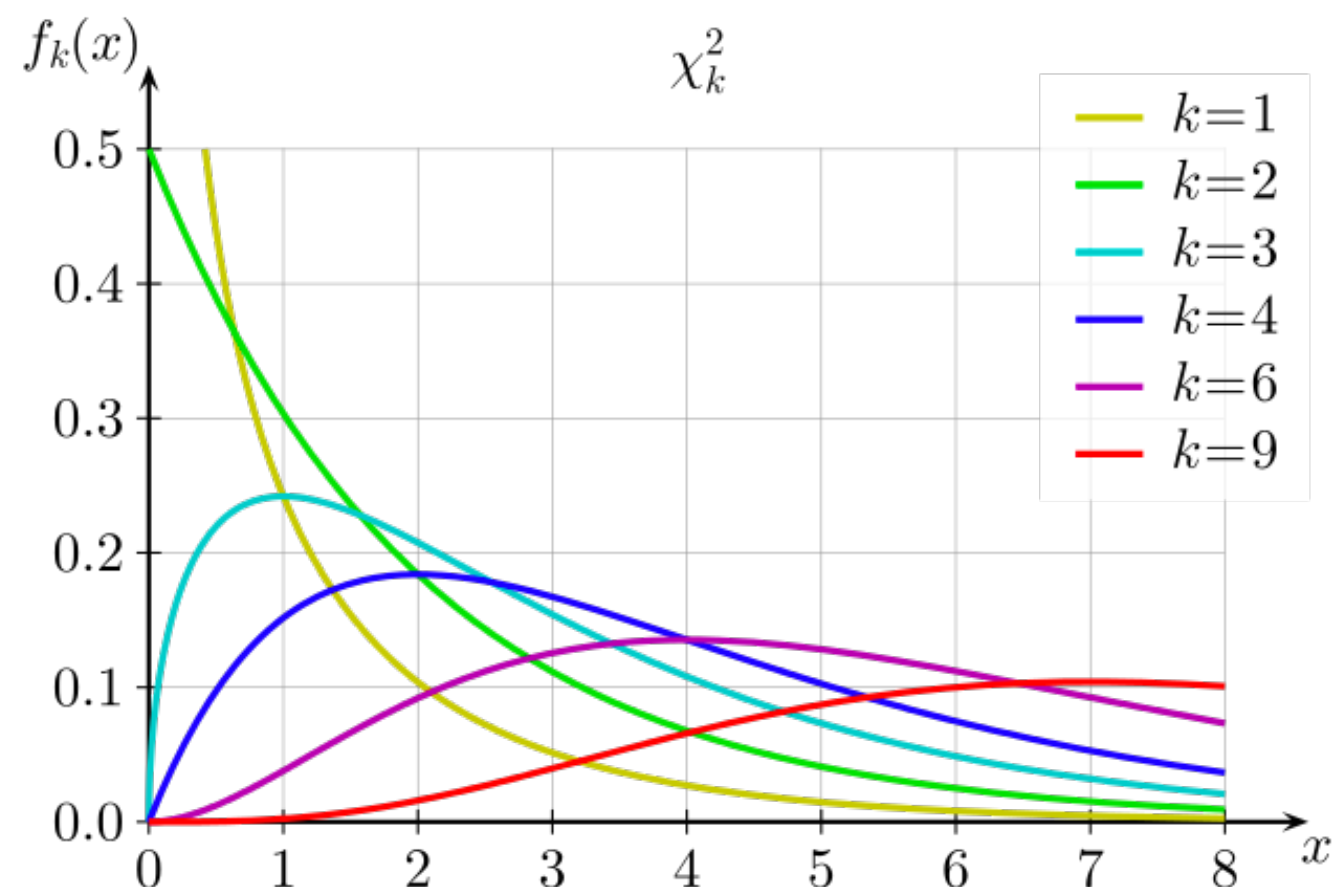
$$\chi^2(x, n) = \sum_{i=1}^N \frac{(x - \mu)^2}{\sigma^2}$$

$$E[\chi^2(x)] \approx N$$

$$E[\chi^2(x)/N] \approx 1$$

$$\text{Var}(\chi^2(x)) = 2N$$

$$\Delta \chi^2(x) = |\chi^2(x) - \chi^2(x \pm \sqrt{2N})|$$



# Understanding

# Important Properties

- Taylor expand in our floated parameter ( $\mu$  in this case)

$$\chi^2(x, n) = \sum_{i=1}^N \frac{(x_i - \mu)^2}{\sigma_i^2} = \underbrace{\sum_{i=2}^N \frac{(x_i - \mu)^2}{\sigma_i^2}}_{\text{red box}} + \underbrace{\frac{(x_1 - \mu)^2}{\sigma_1^2}}_{\text{blue box}}$$

Consider this as a fixed constant



Motion of this point by 1 standard deviation in  $\sigma_1$  causes  $\Delta\chi^2 = 1$  from minimum

Can view the distribution of just  $x_1$  as just a  $\chi^2$  of 1 DOF

# Important Properties

- Taylor expand in our floated parameter ( $\mu$  in this case)

$$\chi^2(x, n) = \sum_{i=1}^N \frac{(x - \mu)^2}{\sigma^2}$$

$$\chi^2(x_i, \mu) = \chi_{min}^2(x_i, \mu_0) + \frac{1}{2} \frac{\partial^2}{\partial \mu^2} \chi_{min}^2(x_i, \mu_0) (\mu - \mu_0)^2$$

Frozen

Varying

$$\Delta \chi^2(x, n) = 1 \rightarrow (\mu \rightarrow \mu \pm \sigma)$$

$$1 = \frac{1}{2} \frac{d}{d\mu^2} \chi^2(\mu_0) \sigma^2$$

# Important Properties

- Taylor expand in our floated parameter ( $\mu$  in this case)

$$\chi^2(x, n) = \sum_{i=1}^N \frac{(x - \mu)^2}{\sigma^2}$$

$$\chi^2(x_i, \mu) = \chi_{min}^2(x_i, \mu_0) + \frac{1}{2} \frac{\partial^2}{\partial \mu^2} \chi_{min}^2(x_i, \mu_0) (\mu - \mu_0)^2$$

Frozen

Varying

$$\Delta\chi^2(x, n) = 1 \rightarrow (\mu \rightarrow \mu \pm \sigma) \quad \textbf{Wilk's Theorem}$$

$$1 = \frac{1}{2} \frac{d}{d\mu^2} \chi^2(\mu_0) \sigma^2 \rightarrow \boxed{\sigma^2 = \frac{2}{\frac{d}{d\mu^2} \chi^2(\mu_0)}}$$

# An Example

Recall that if we vary take the average over N

Our uncertainty on the mean goes as

$$\sigma_{\mu} = \sigma \sqrt{\frac{1}{N}}$$



# An Example

$$\Delta\chi^2 = 1 = \sum_{i=1}^N \frac{(x - \mu_0 + \sigma_\mu)^2}{\sigma^2} - \sum_{i=1}^N \frac{(x - \mu_0)^2}{\sigma^2}$$

$$1 \approx \sum_{i=1}^N \frac{(x - \mu_0 + \sigma_\mu)^2}{\sigma^2} - \sum_{i=1}^N \frac{(x - \mu_0)^2}{\sigma^2}$$

$$1 = \frac{1}{\sigma^2} \sum_{i=1}^N (x - \mu_0 + \sigma_\mu)^2 - (x - \mu_0)^2$$

$$1 = \frac{1}{\sigma^2} \sum_{i=1}^N \sigma_\mu^2 + 2\sigma_\mu(x - \mu_0)$$

$$1 = \frac{N\sigma_\mu^2}{\sigma^2}$$

$$\sigma_\mu^2 = \frac{\sigma^2}{N}$$



For a poisson distribution we recover the variance per bin

# Important Properties

- Taylor expand in our floated parameter ( $\mu$  in this case)

$$\chi^2(x_i, \mu) = \chi_{min}^2(x_i, \mu_0) + \frac{1}{2} \frac{\partial^2}{\partial \mu^2} \chi_{min}^2(x_i, \mu_0) (\mu - \mu_0)^2$$

$\chi^2$  distribution of 1 degree of freedom  
 $V[\chi^2(x)] = 1$

$$\Delta \chi^2 = 2 \Delta \log L = 1$$

For one degree of freedom

# Important Properties

- Taylor expand in our floated parameter ( $\mu$  in this case)

$$\chi^2(x_i, \mu) = \chi_{min}^2(x_i, \mu_0) + \frac{1}{2} \frac{\partial^2}{\partial \mu^2} \chi_{min}^2(x_i, \mu_0) (\mu - \mu_0)^2$$

$$\frac{1}{2} \frac{d^2 \chi^2}{d\mu^2} \rightarrow \frac{1}{\sigma^2}$$

$$\frac{\partial^2 \chi^2}{\partial \theta^2} = \frac{2}{\sigma_\theta^2}$$

For any floated parameter  
uncertainty of that parameter is  
given by the 2nd derivative of  $\chi^2$

This is known as Wilk's Theorem

$$\sigma_\theta^2 = \left( \frac{\partial^2 \log L}{\partial \theta^2} \right)^{-1}$$

# Important Properties

- Taylor expand in our floated parameter ( $\mu$  in this case)

$$\chi^2(x_i, \mu) = \chi_{min}^2(x_i, \mu_0) + \frac{1}{2} \frac{\partial^2}{\partial \mu^2} \chi_{min}^2(x_i, \mu_0) (\mu - \mu_0)^2$$

$$\frac{1}{2} \frac{d^2 \chi^2}{d\mu^2} \rightarrow \frac{1}{\sigma^2}$$

$\chi^2$  distribution of 1 degree of freedom  
 $V[\chi^2(x)] = 1$

$$\frac{\partial^2 \chi^2}{\partial \theta^2} = \frac{2}{\sigma_\theta^2}$$

$$\Delta \chi^2 = 2 \Delta \log L = 1$$

For one degree of freedom

For any floated parameter  
 uncertainty of that parameter is  
 given by the 2nd derivative of  $\chi^2$

This is known as Wilk's Theorem

$$\sigma_\theta^2 = \left( \frac{\partial^2 \log L}{\partial \theta^2} \right)^{-1}$$

# Multiple Dimensions

- For N variables the expansion is the same

$$\chi^2(x_i, \vec{\theta}) = \chi_{min}^2(x_i, \vec{\theta}) + \frac{1}{2}(\theta_i - \theta_0)^T \frac{\partial^2}{\partial \theta_i \partial \theta_j} \chi_{min}^2(x_i, \vec{\theta}_0)(\theta_j - \theta_0)$$

$\chi^2$  distribution of 1 degree of freedom  
 $V[\chi^2(x)] = 1$

$$\Delta \chi^2 = 2 \Delta \log L = 1$$

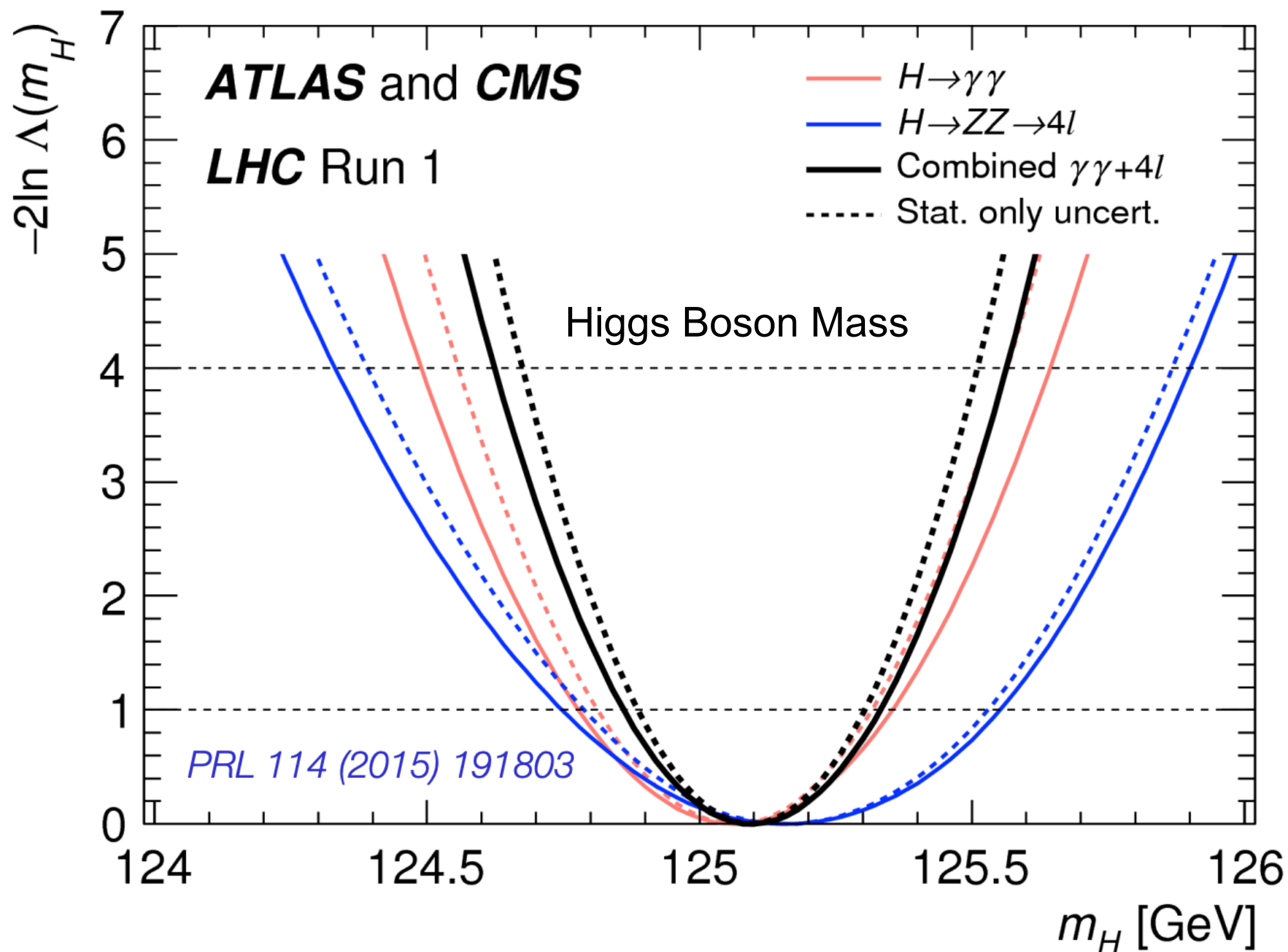
For one degree of freedom

Hessian of  
 the  $\chi^2$  distribution

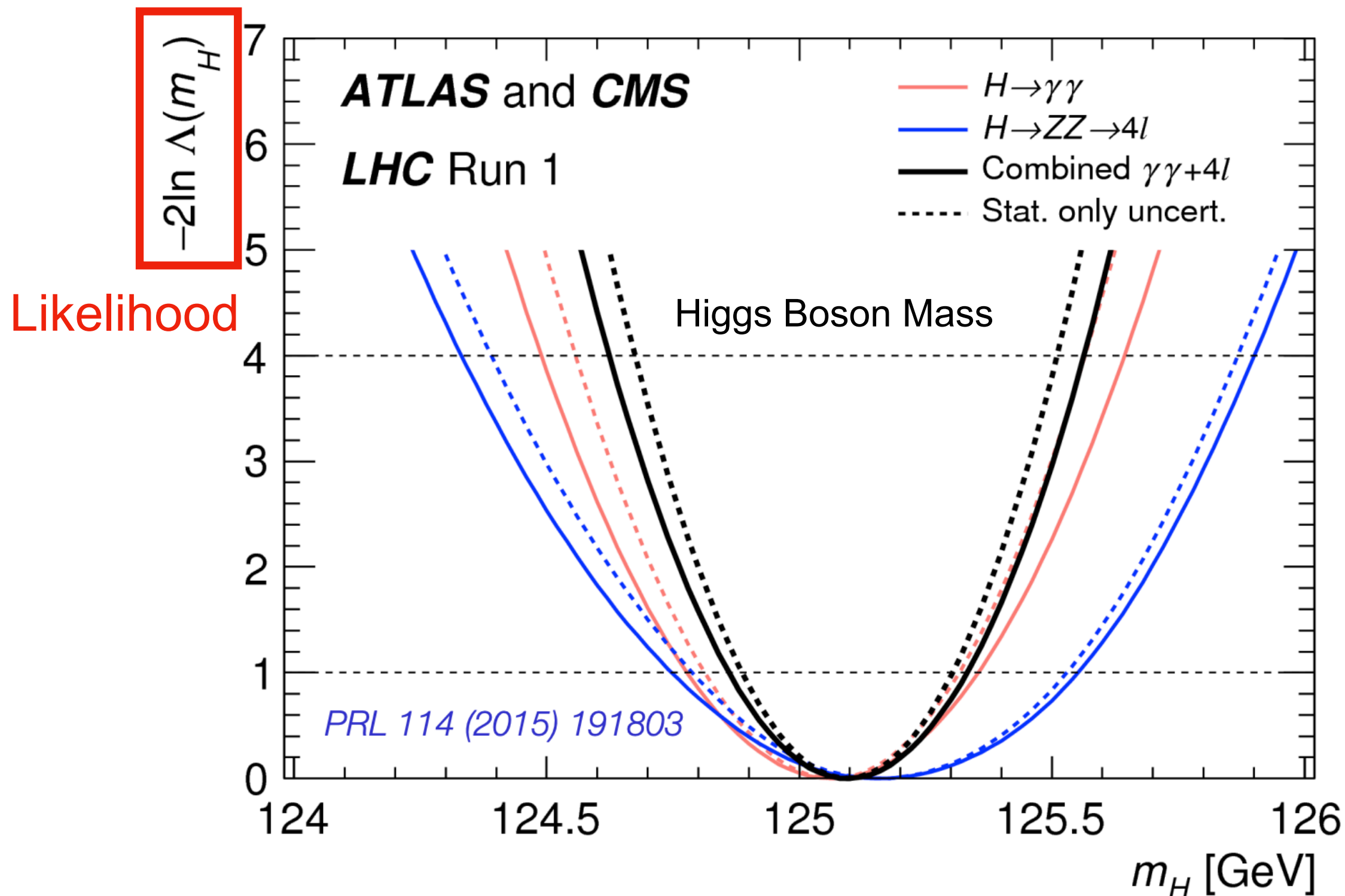
This is known as Wilk's Theorem

$$\sigma_{ij}^2 = \left( \frac{\partial^2 \log L}{\partial \theta_i \partial \theta_j} \right)^{-1}$$

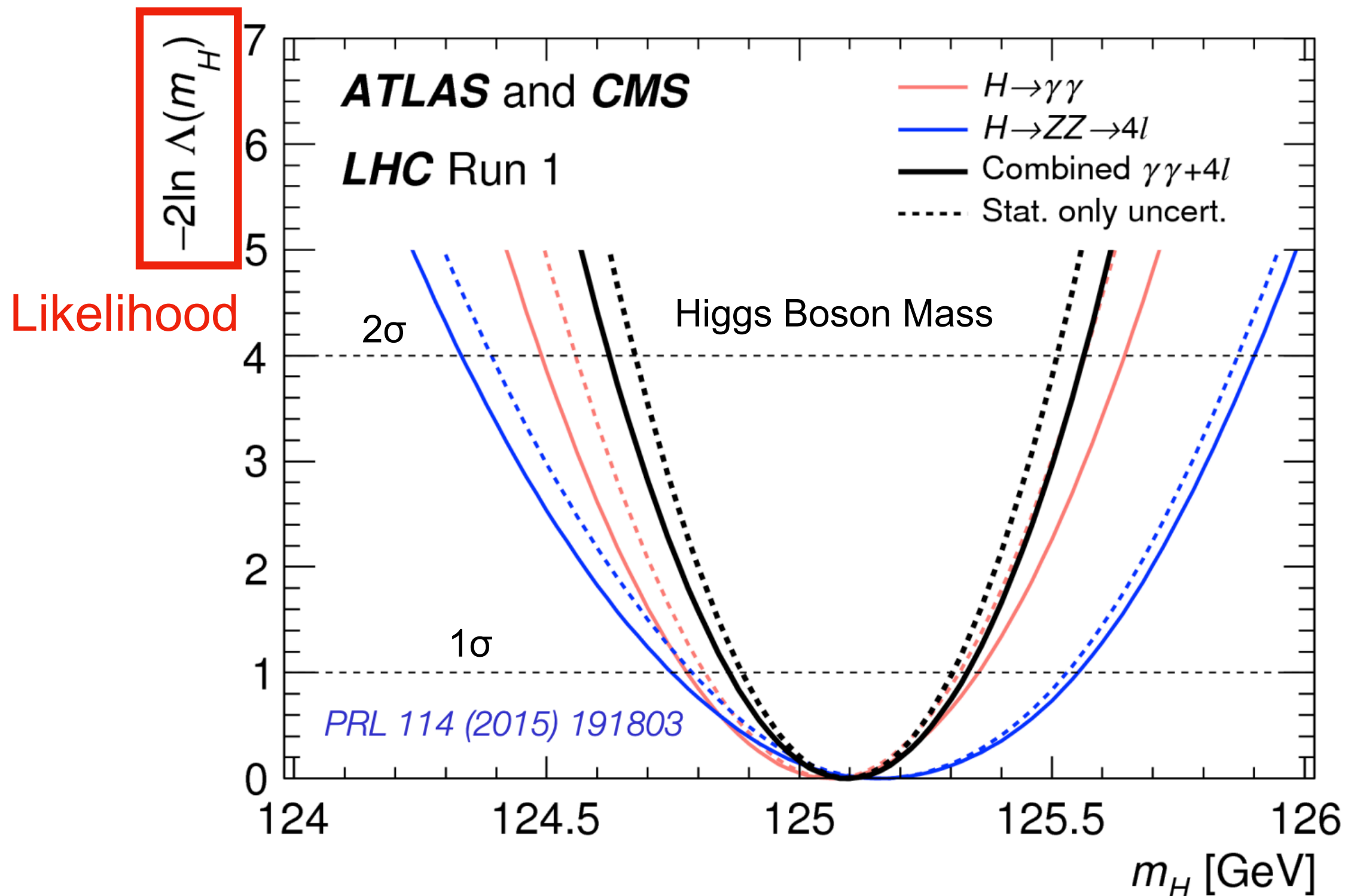
# Making a Measurement



# Making a Measurement



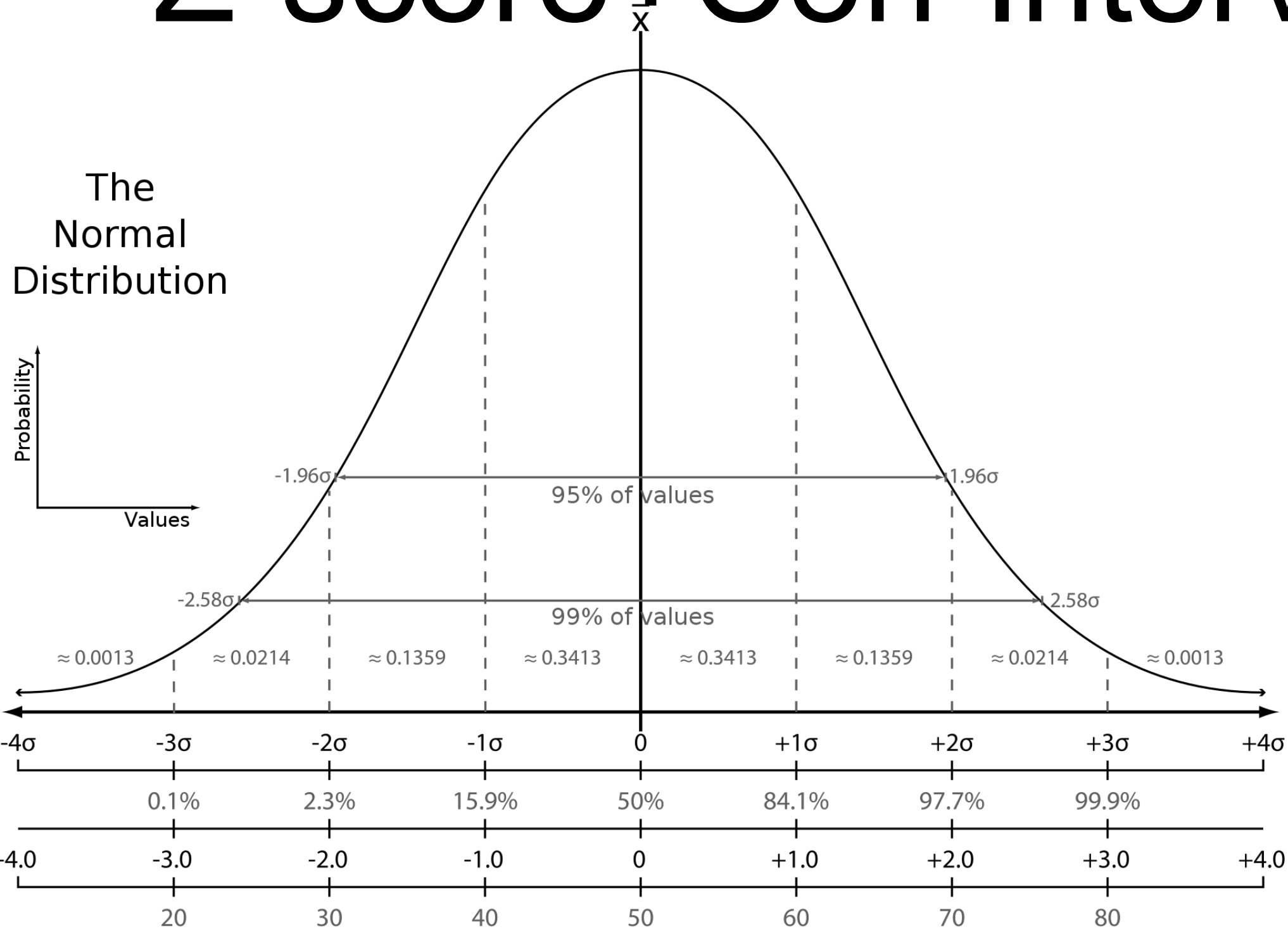
# Making a Measurement



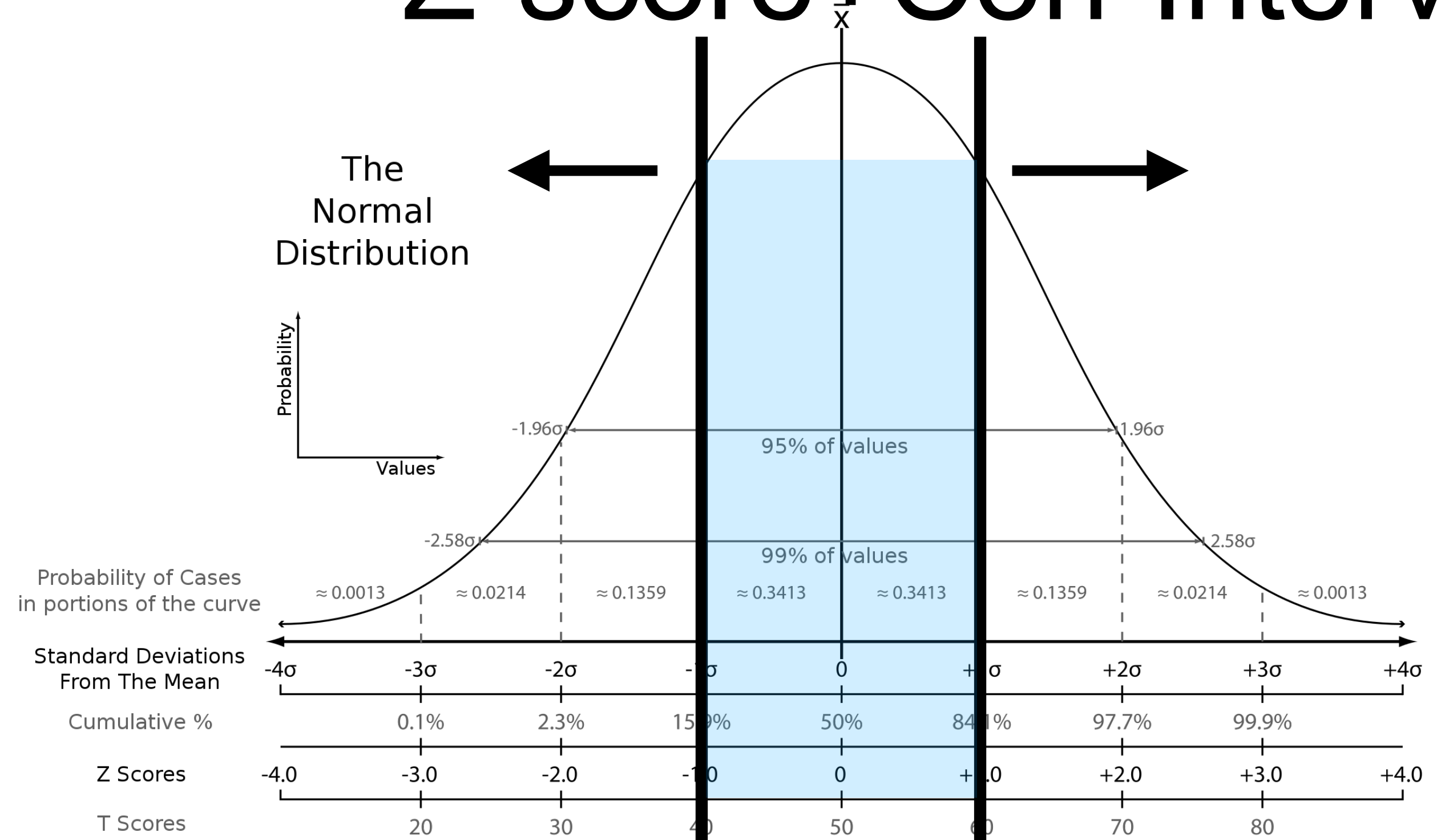


# Confidence Intervals

# Z-score+Con-Interval

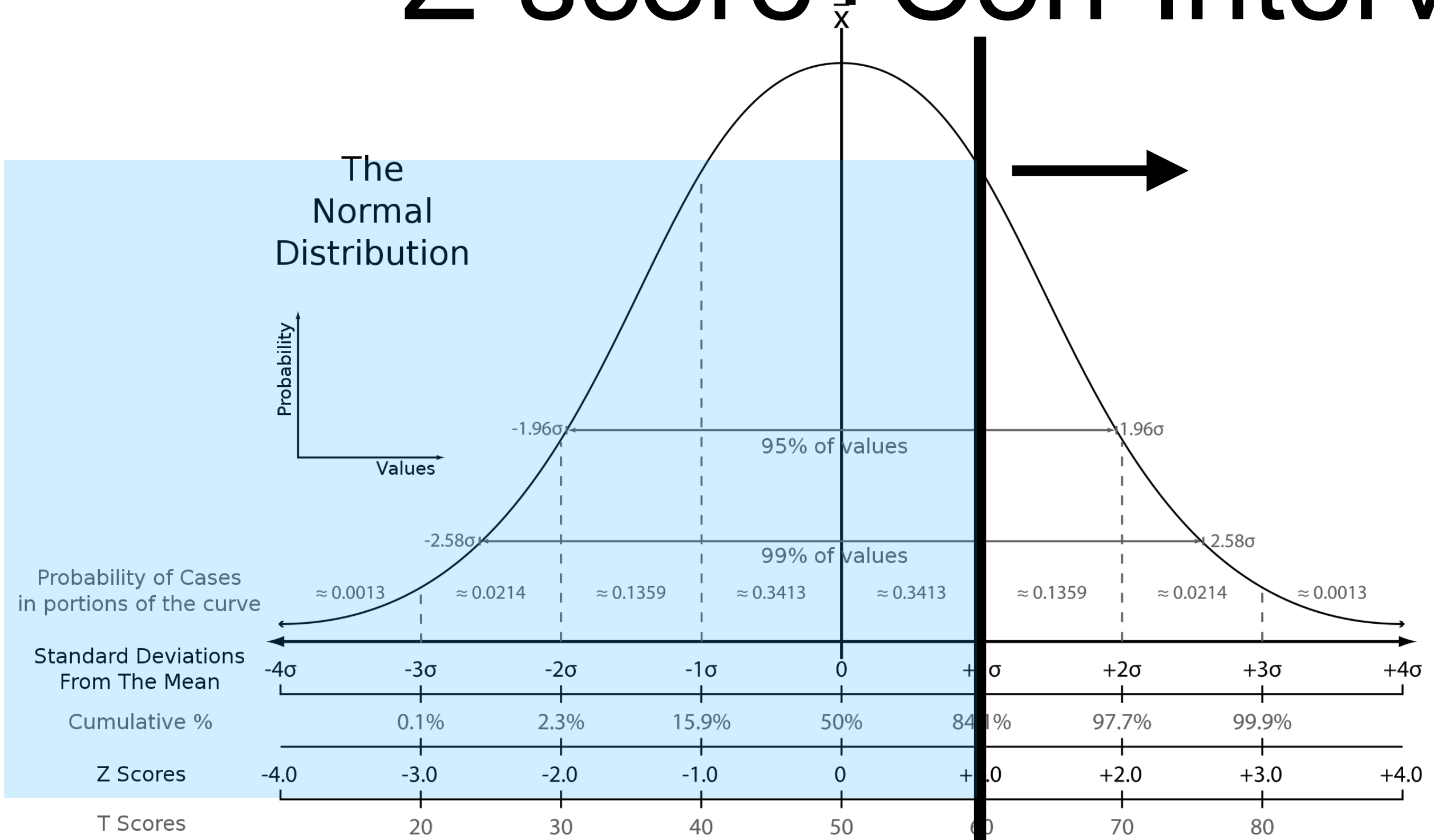


# Z-score+Con-Interval



Z-score of 1: 67% chance of being within 1 standard deviation

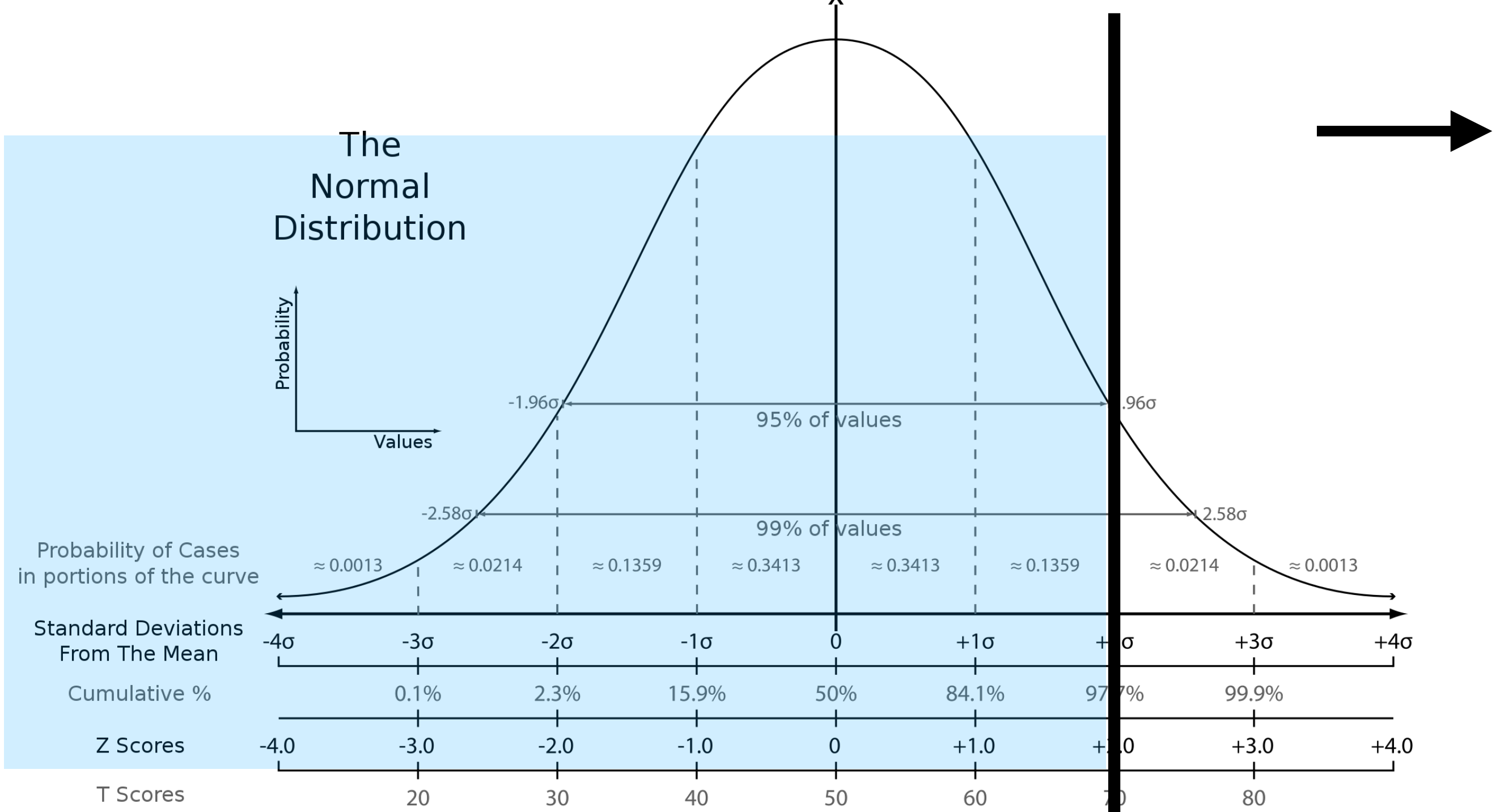
# Z-score+Con-Interval



Single sided Z-score of 1: 84% chance of being above 1 standard deviation



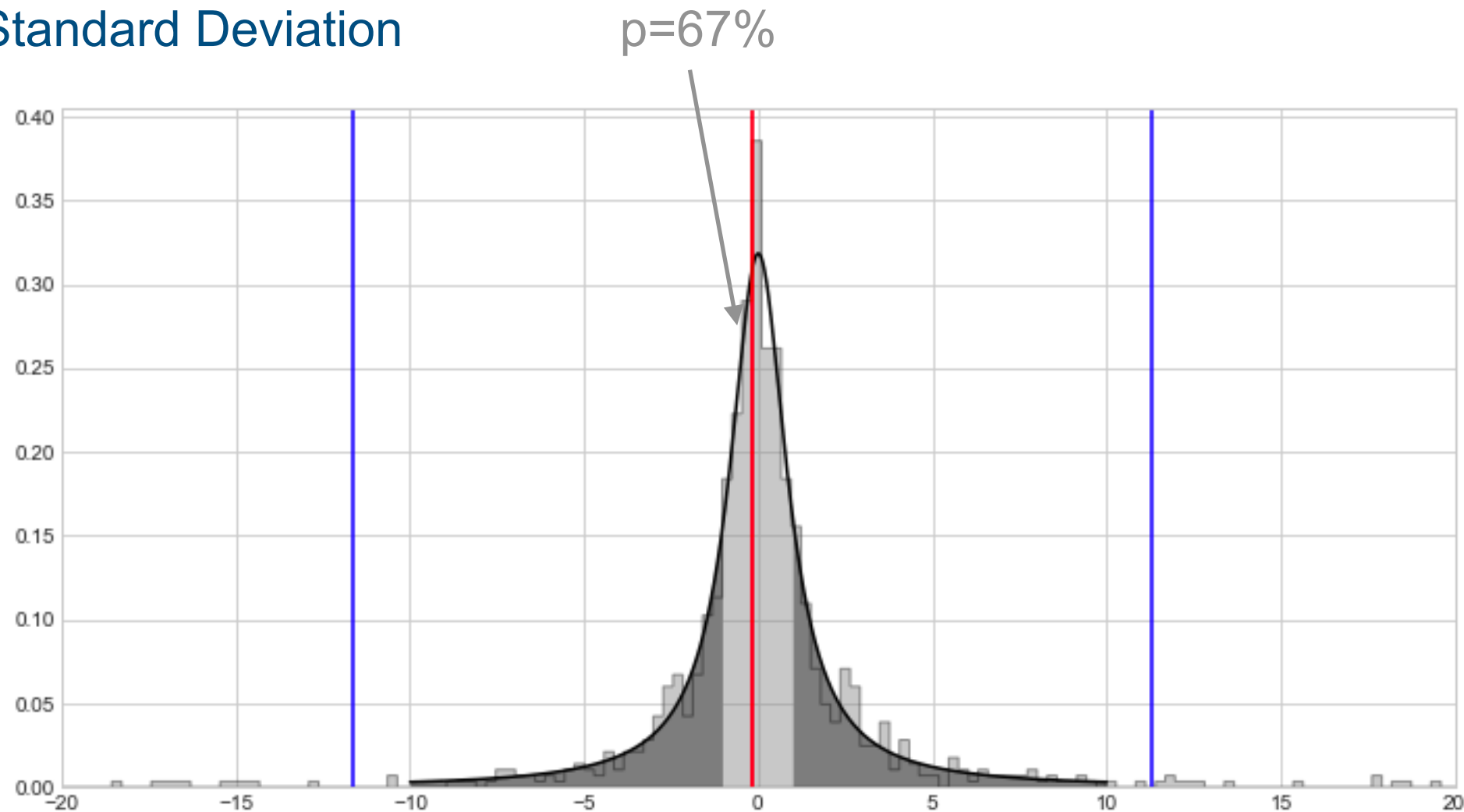
# Z-score+Con-Interval



Single sided Z-score of 2: 97% chance of being above 1 standard deviation

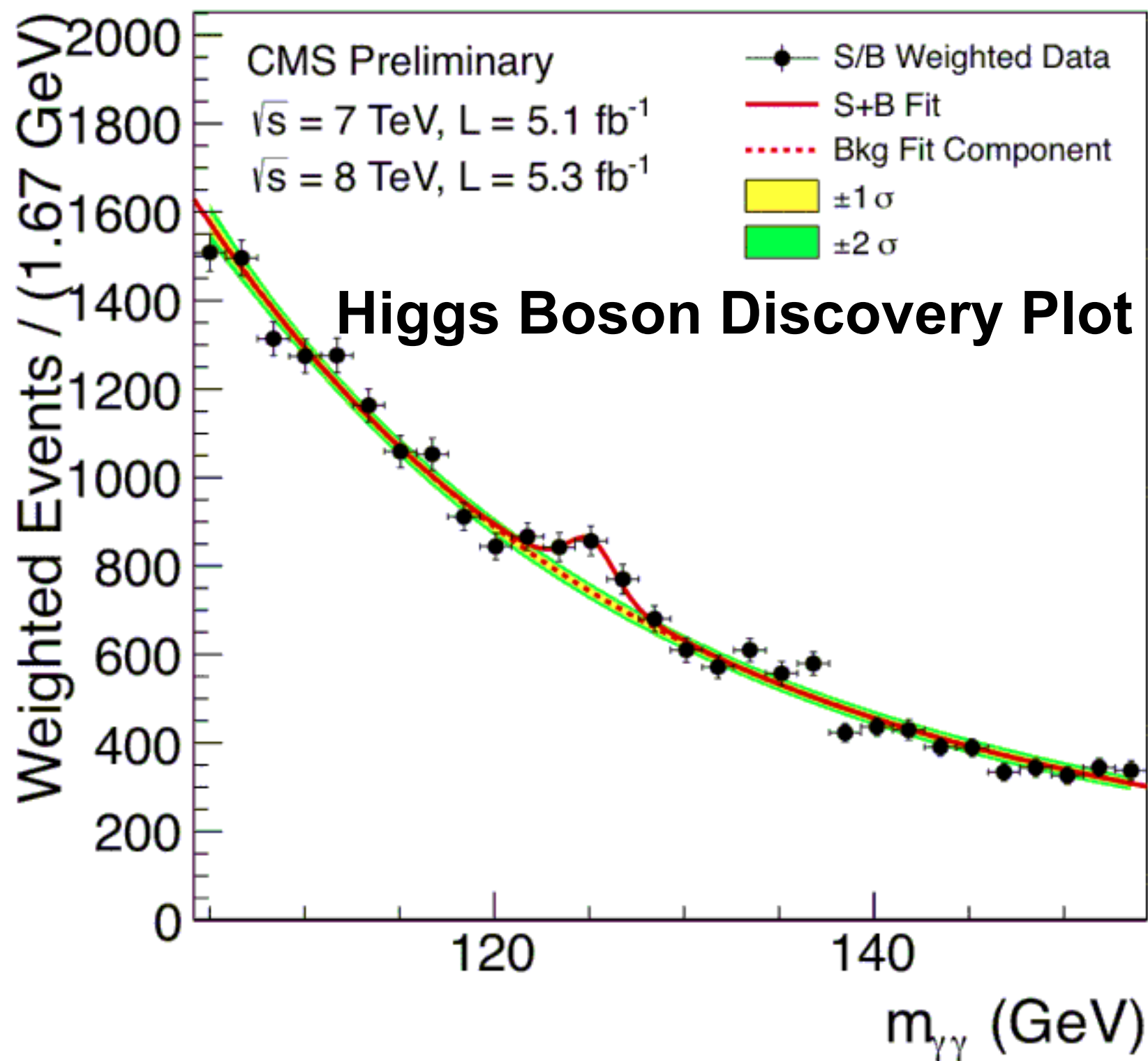
# Z-score+Con-Interval

Standard Deviation



- Z score works on any system
- However standard deviation does not necessarily reflect the z-score

# Confidence Plots





# The rules of significance

- How significant is our measurement? (High Energy physics rules)
- 3 sigma is considered “Evidence”
- 5 sigma is considered “Discovery”

<https://understandinguncertainty.org/explaining-5-sigma-higgs-how-well-did-they-do>



The screenshot shows the homepage of the 'Understanding Uncertainty' website. The header features a yellow 'uu' logo and a navigation bar with links: Home, Blog, Articles, Videos, Animations, Guest Articles, Links, and About Us. The main content area displays a blog post titled 'Explaining 5-sigma for the Higgs: how well did they do?' by david, dated 08/07/2012. The post includes a warning that the content is for statistical pedants and begins with a recap of Higgs results. On the right side, there is a search bar, a dropdown menu for featured content, and a main menu link.

uu

## Understanding Uncertainty

Home Blog Articles Videos Animations Guest Articles Links About Us

Home » Blogs » david's blog

### Explaining 5-sigma for the Higgs: how well did they do?

Submitted by david on Sun, 08/07/2012 - 1:17pm

Warning, this is for statistical pedants only.

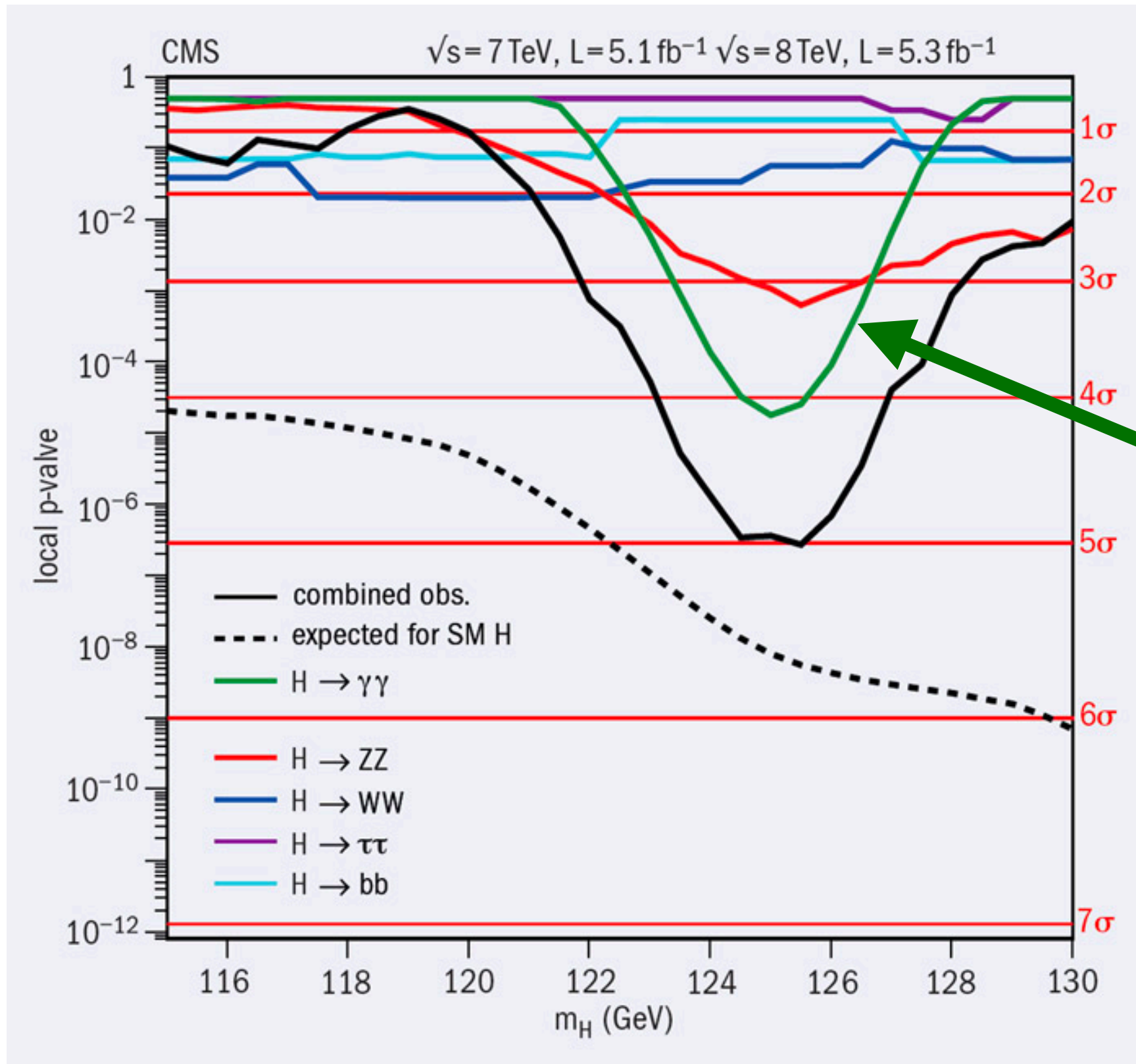
To recap, the results on the Higgs are communicated in terms of the numbers of

Search

- Featured Content -

Main menu

# Confidence Plots

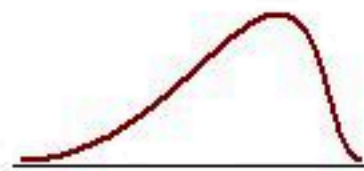


# Moments

$$\mu_n = m^n(x) = E[x^n p(x)] = \int_{-\infty}^{\infty} x^n p(x) dx$$

## Skewness

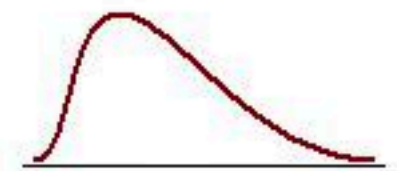
The coefficient of Skewness is a measure for the degree of symmetry in the variable distribution.



Negatively skewed distribution  
or Skewed to the left  
Skewness < 0



Normal distribution  
Symmetrical  
Skewness = 0

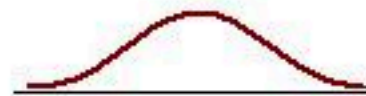


Positively skewed distribution  
or Skewed to the right  
Skewness > 0

- Moments are a way to characterize the function
- n=1 is mean
- n=2 is variance
- n=3 is Skew
- n=4 is kurtosis

## Kurtosis

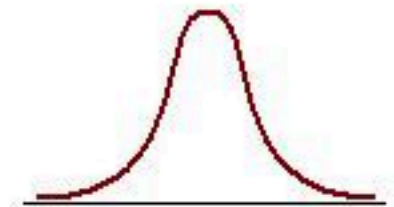
The coefficient of Kurtosis is a measure for the degree of peakedness/flatness in the variable distribution.



Platykurtic distribution  
Low degree of peakedness  
Kurtosis < 0

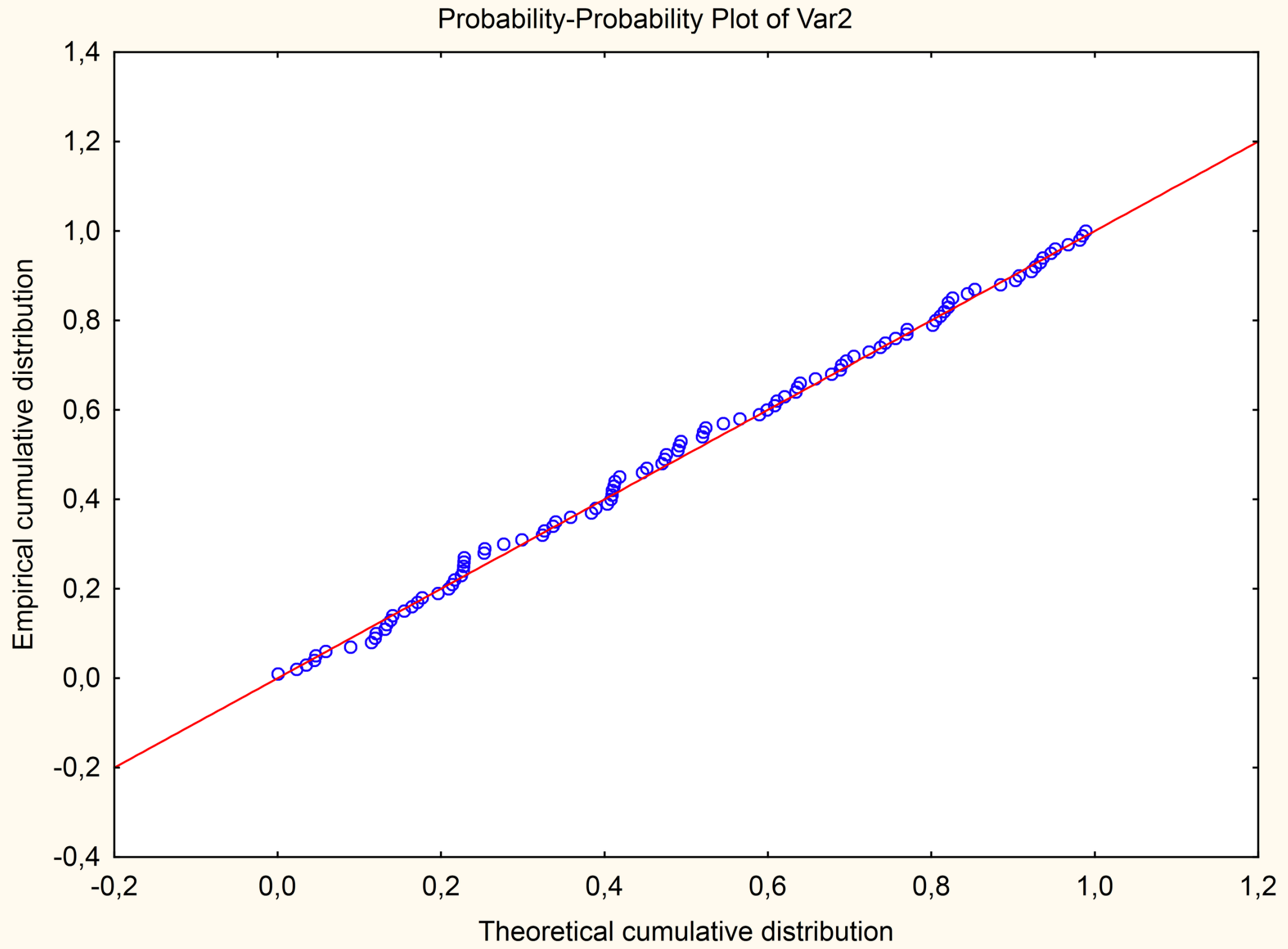


Normal distribution  
Mesokurtic distribution  
Kurtosis = 0



Leptokurtic distribution  
High degree of peakedness  
Kurtosis > 0

# Dn\_plot





# Our final expansion Plot

