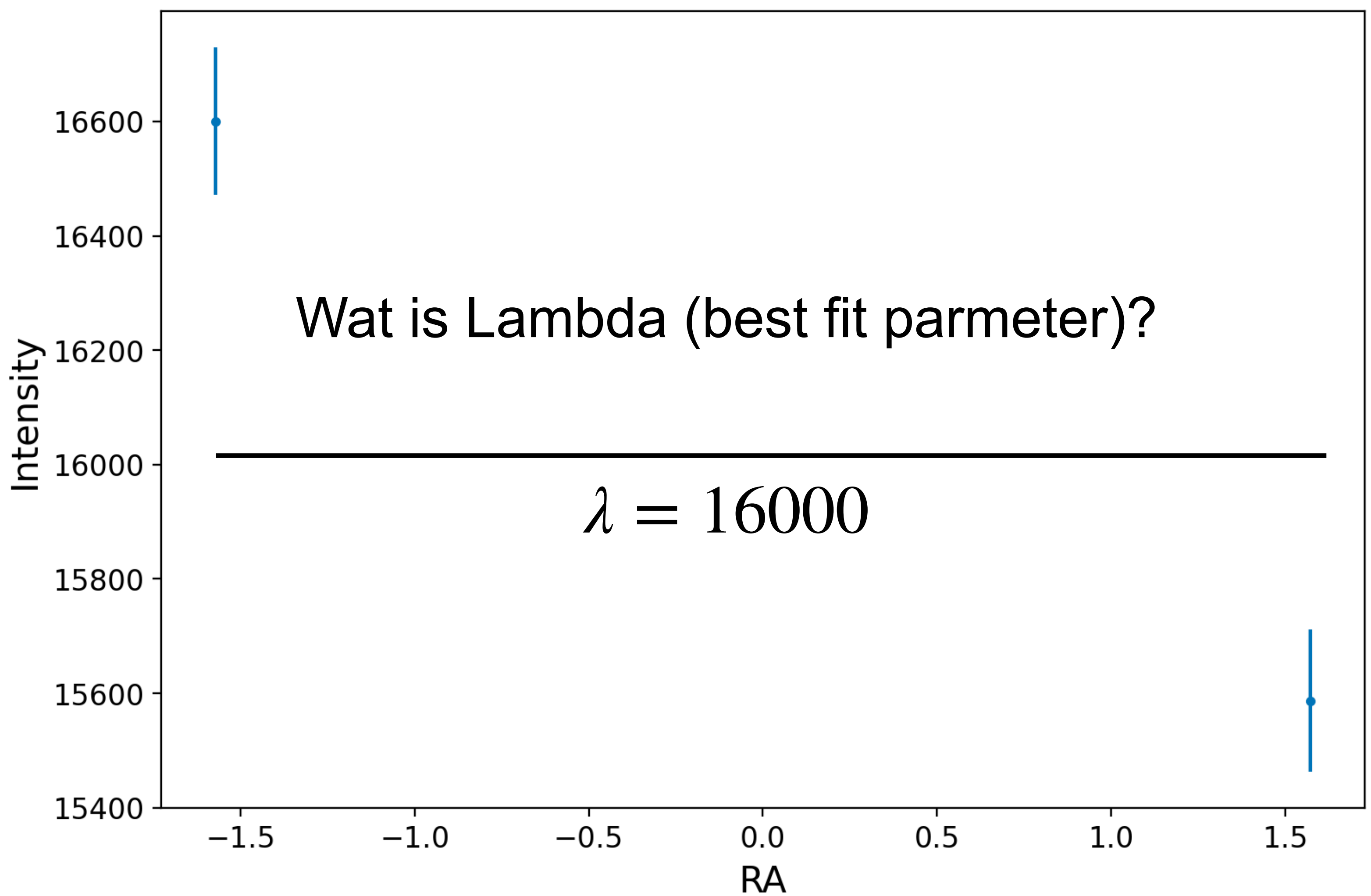
Lecture 6:Confidence

Recap

Likelihood fitting

- Given a probability distribution with a free parameter
    - We wanted to fit that free parameter



Likelihood fitting

$$\mathcal{L}(x|\lambda) = \prod_{i=1}^{N} p(x_i|\lambda)$$

$$\mathcal{L}(x|\lambda) = \prod_{i=1}^{N} \frac{\lambda^{x_i}}{x_i!} e^{-\lambda}$$

Numerically

$$\text{-}\ \log(\mathcal{L}(x|\lambda)) = \sum_{i-1}^{N} x_i \log(\lambda) - \log(x_i!) - \lambda$$

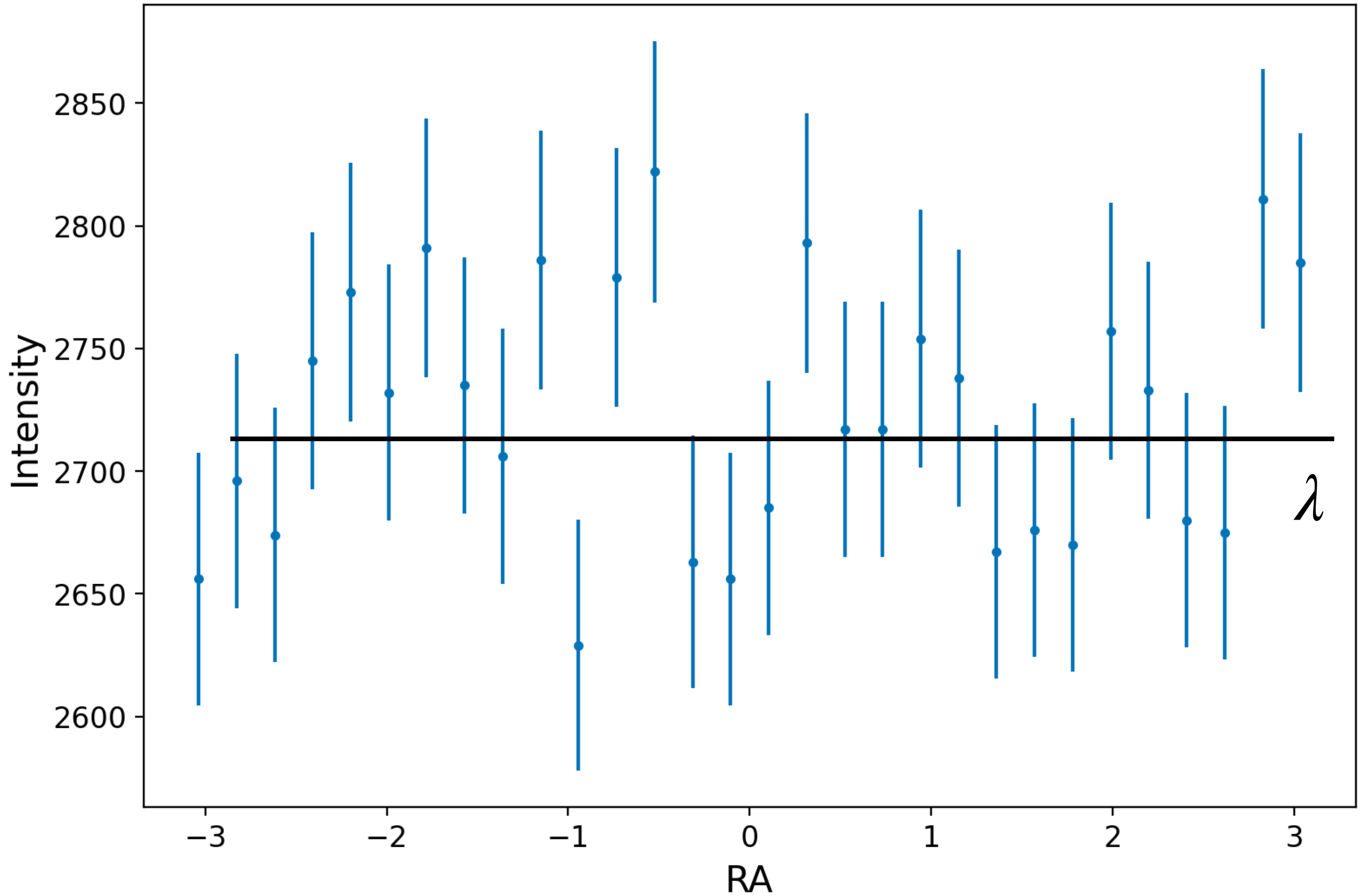
$$-\log(\mathcal{L}(x|\lambda)) = \sum_{i-1}^{N} -\frac{1}{2}\log(2\pi\lambda) + \frac{(x_i - \lambda)^2}{2\lambda}$$

$$\frac{d}{d\lambda}\log(\mathcal{L}(x|\lambda)) = 0 = \sum_{i-1}^{N} \frac{x_i}{\lambda} - 1$$

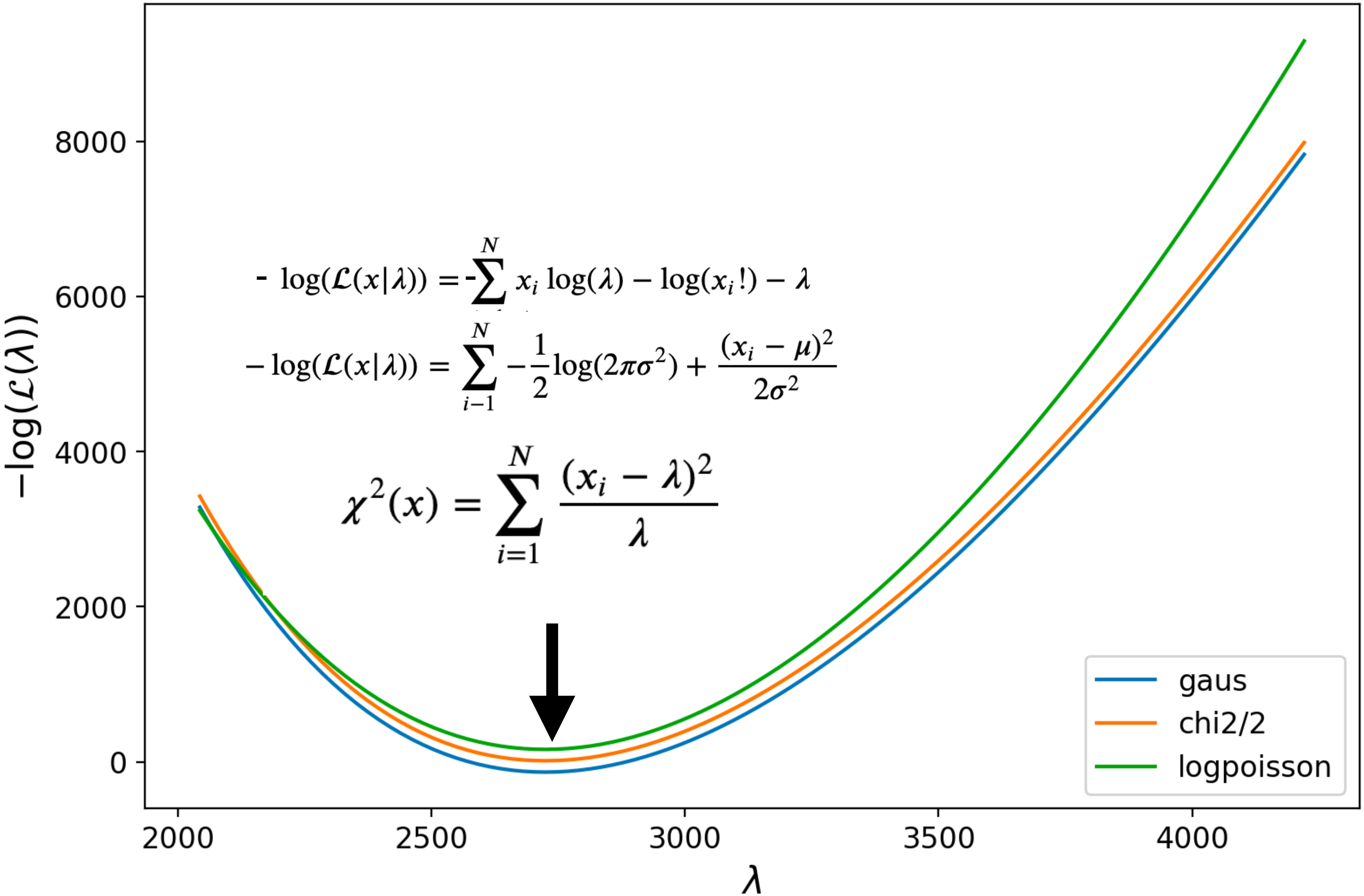
$$N\lambda = \sum_{i-1}^{N} x_i$$

$$\lambda = \bar{x}$$

Analytically



Likelihood fitting



$$-\log(\mathcal{L}(x|\lambda)) = -\sum^{N} x_i \log(\lambda) - \log(x_i!) - \lambda$$

$$-\log(\mathcal{L}(x|\lambda)) = \sum_{i-1}^{N} -\frac{1}{2}\log(2\pi\sigma^2) + \frac{(x_i-\mu)^2}{2\sigma^2}$$

$$\chi^2(x) = \sum_{i=1}^{N} \frac{(x_i-\lambda)^2}{\lambda}$$

χ^2 distribution

- χ^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

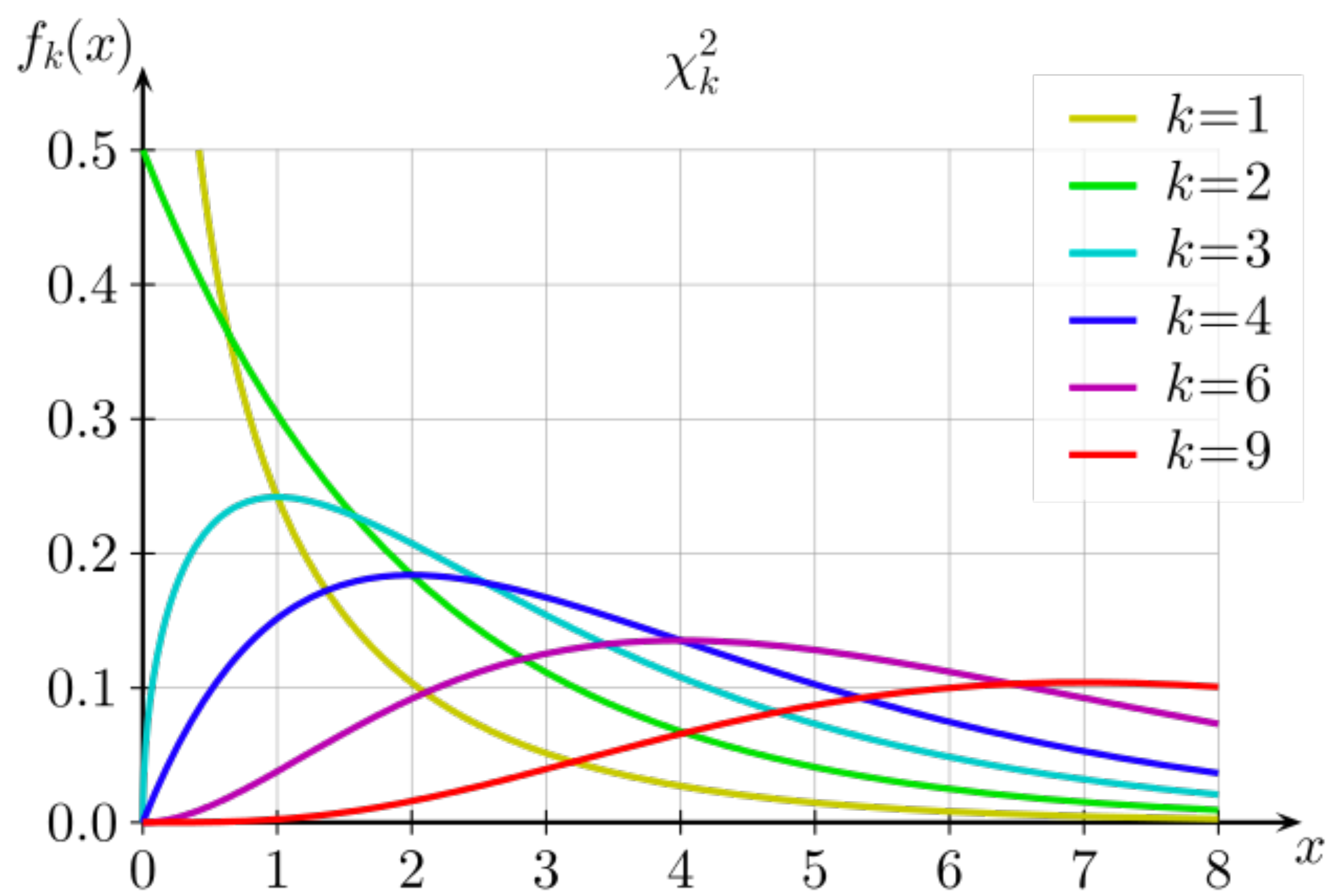
$$\chi^2(x,n) = \sum_{i=1}^{N} \frac{(x-\mu)^2}{\sigma^2}$$

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

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

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

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



Understanding

Important Properties

- Taylor expand in our floated parameter (μ in this case)

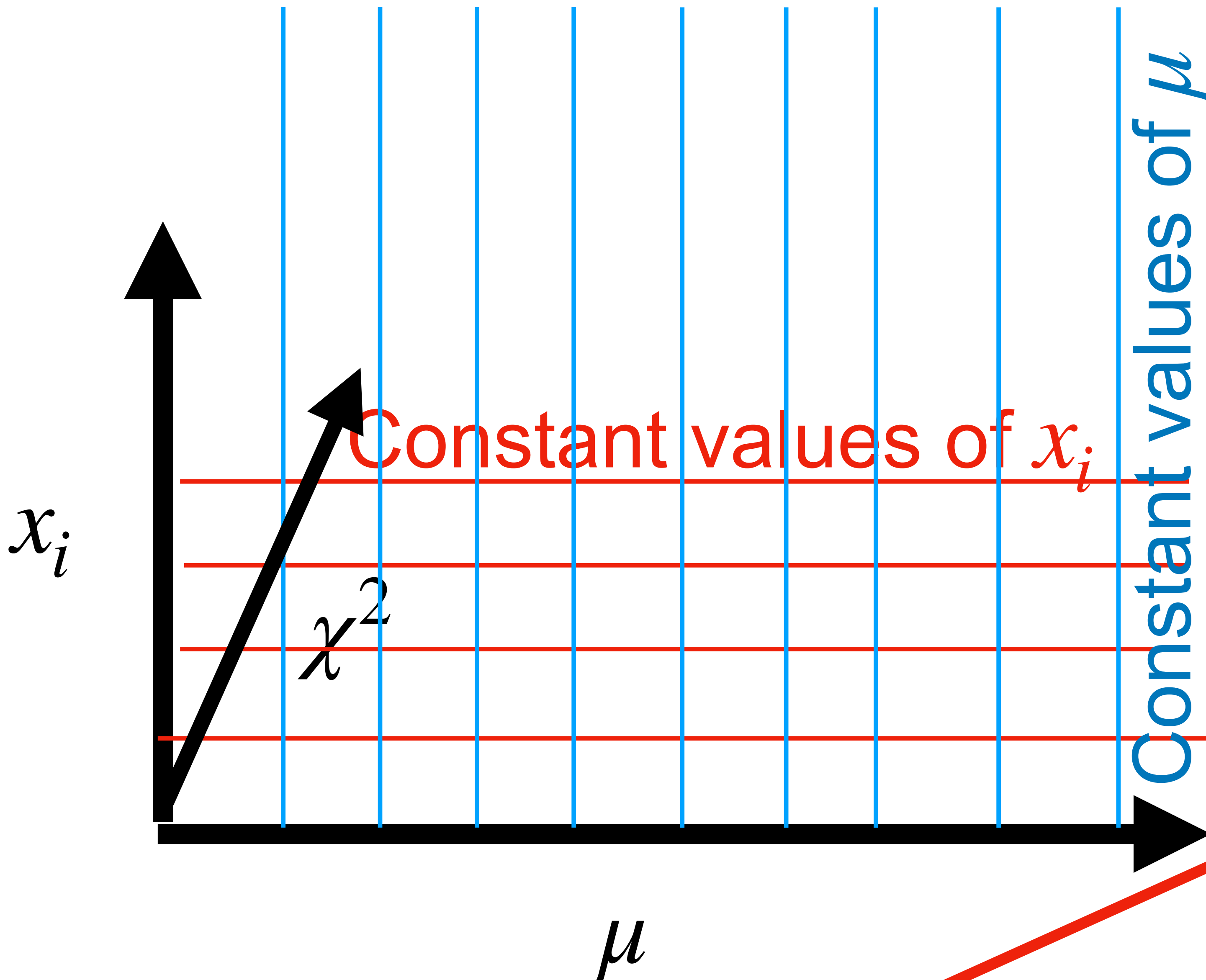
$$\chi^2(x,n) = \sum_{i=1}^{N} \frac{(x_i-\mu)^2}{\sigma_i^2} = \boxed{\sum_{i=2}^{N} \frac{(x_i-\mu)^2}{\sigma_i^2}} + \boxed{\frac{(x_1-\mu)^2}{\sigma_1^2}}$$

Consider this as a fixed constant

Motion of this point by 1 standard deviation in σ_1 causes $\Delta\chi^2 = 1$ from minimum

Can view the distribution of just x_1 as just a χ^2 of 1 DOF

Visualization



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

A function of x_i and μ

Visualization

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

Chi-2 of N-degrees of freedom

$$\chi^2(x,1) = \frac{(\mu - \mu_0)^2}{\sigma_\mu^2} + C_{mu}$$

Chi-2 of 1-degree of freedom

Visualization

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

Chi-2 of N-degrees of freedom

$$\chi^2(x,1) = \frac{(\mu - \mu_0)^2}{\sigma_\mu^2} + C_{mu}$$

Chi-2 of 1-degree of freedom

These are the same formula

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

Important Properties

- Taylor expand in our floated parameter (μ in this case)

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

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

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 (μ in this case)

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

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

Frozen Varying

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

Wilk's Theorem

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

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

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

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

$$1 = \sum_{i=1}^{N} \frac{1}{\sigma_i^2}(x_i - \mu_0 + \sigma_\mu)^2 - (x_o - \mu_0)^2$$

$$1 = \sum_{i=1}^{N} \frac{1}{\sigma_i^2}\left(\sigma_\mu^2 + 2\sigma_\mu(x_i - \mu_0)\right)$$

$$1 = \sum_{i=1}^{N} \frac{1}{\sigma_i^2}\left(\sigma_\mu^2\right)$$

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

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

$$\sigma_\mu^2 = \frac{\mu_0}{N}$$

For a poisson distribution we recover the variance per bin

Important Properties

- Taylor expand in our floated parameter (μ in this case)

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

χ^2 distribution of 1 degree of freedom

V[$\chi^2(x)$]=1

$$\boxed{\Delta\chi^2 = 2\Delta \log \mathrm{L} = 1}$$

For one degree of freedom

Important Properties

- Taylor expand in our floated parameter (μ in this case)

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

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

$$\boxed{\frac{\partial^2\chi^2}{\partial\theta^2} = \frac{2}{\sigma^2_\theta}}$$

For any floated parameter uncertainty of that parameter is given by the 2nd derivative of χ^2

This is known as Wilk's Theorem

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

Important Properties

- Taylor expand in our floated parameter (μ in this case)

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

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

χ^2 distribution of 1 degree of freedom

V[$\chi^2(x)$]=1

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

$$\boxed{\Delta\chi^2 = 2\Delta\log \mathrm{L} = 1}$$

For one degree of freedom

For any floated parameter uncertainty of that parameter is given by the 2nd derivative of χ^2

This is known as Wilk's Theorem

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

Multiple Dimensions

- For N variables the expansion is the same

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

χ^2 distribution of 1 degree of freedom

V[$\chi^2(x)$]=1

Hessian of
the χ^2 distribution

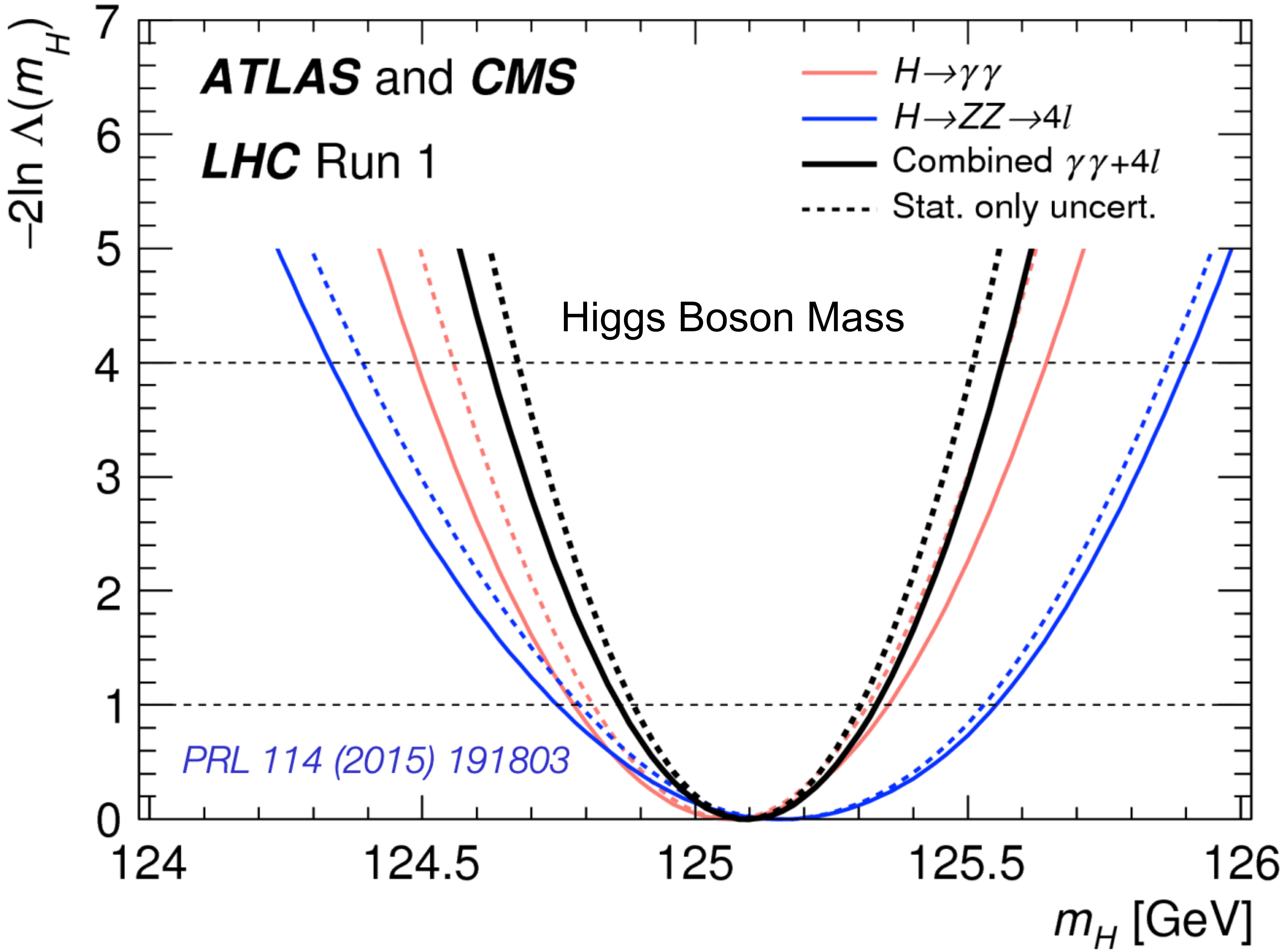
$$\Delta\chi^2 = 2\Delta\log\mathrm{L} = 1$$

For one degree of freedom

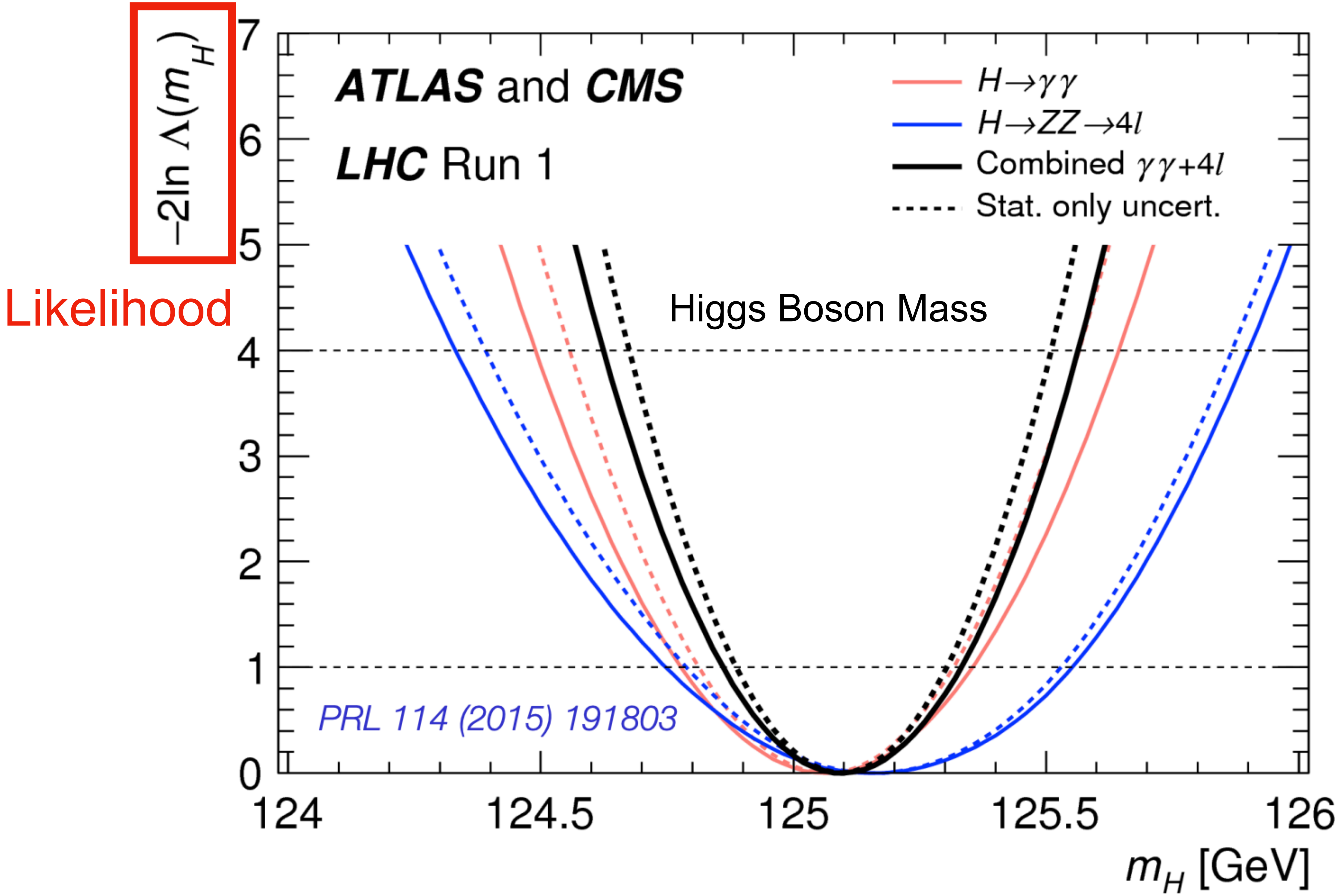
This is known as Wilk's Theorem

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

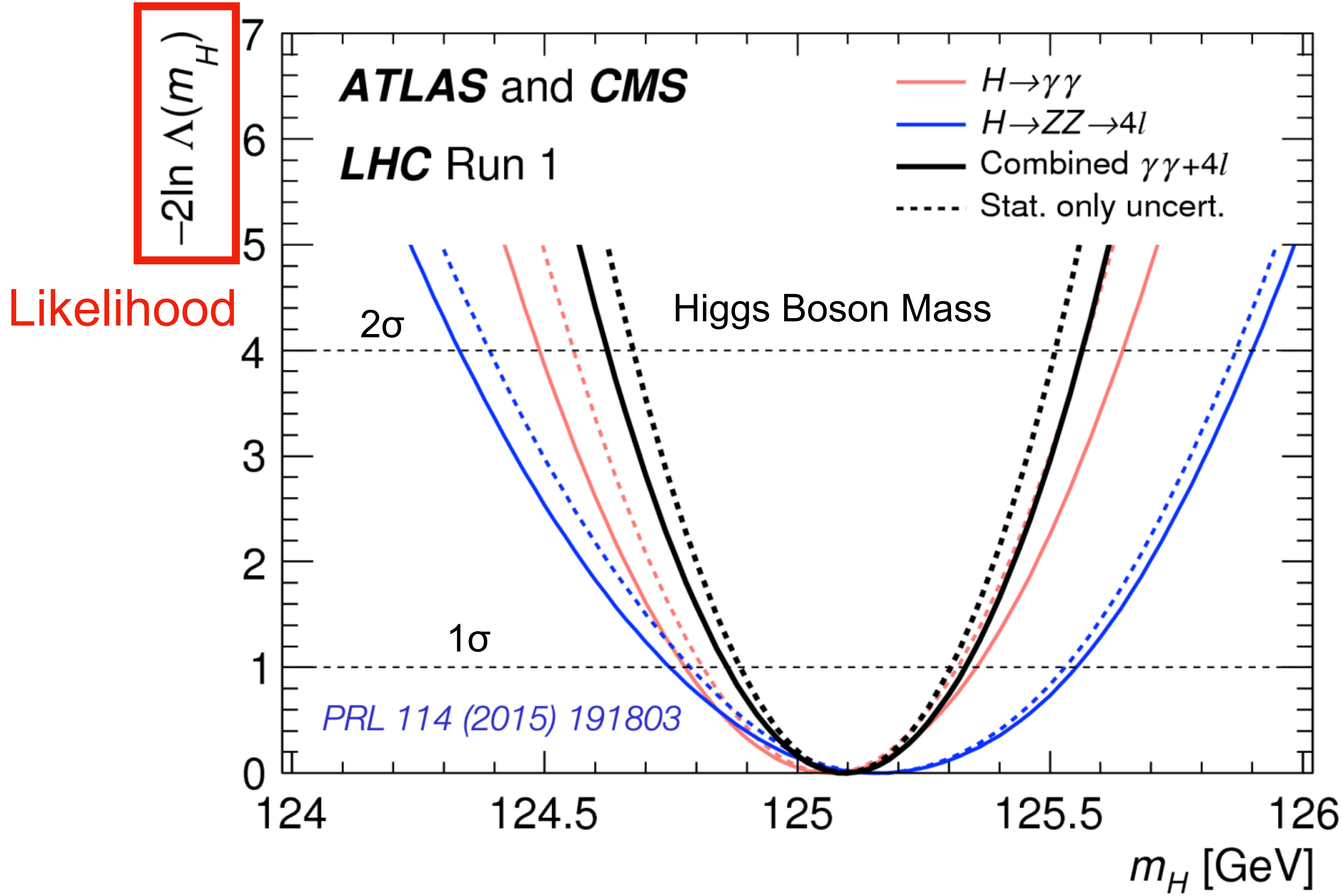
Making a Measurement



Making a Measurement

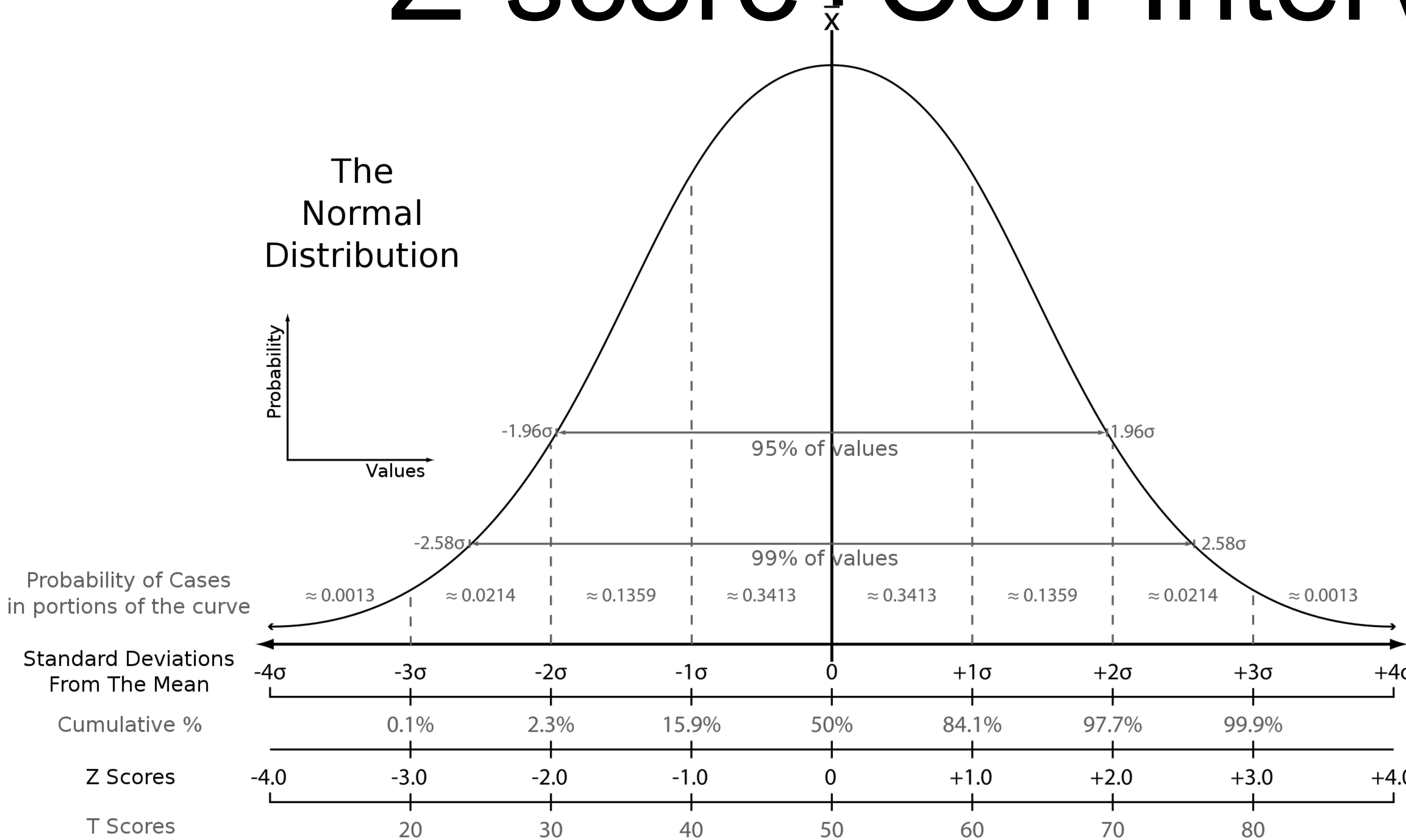


Making a Measurement

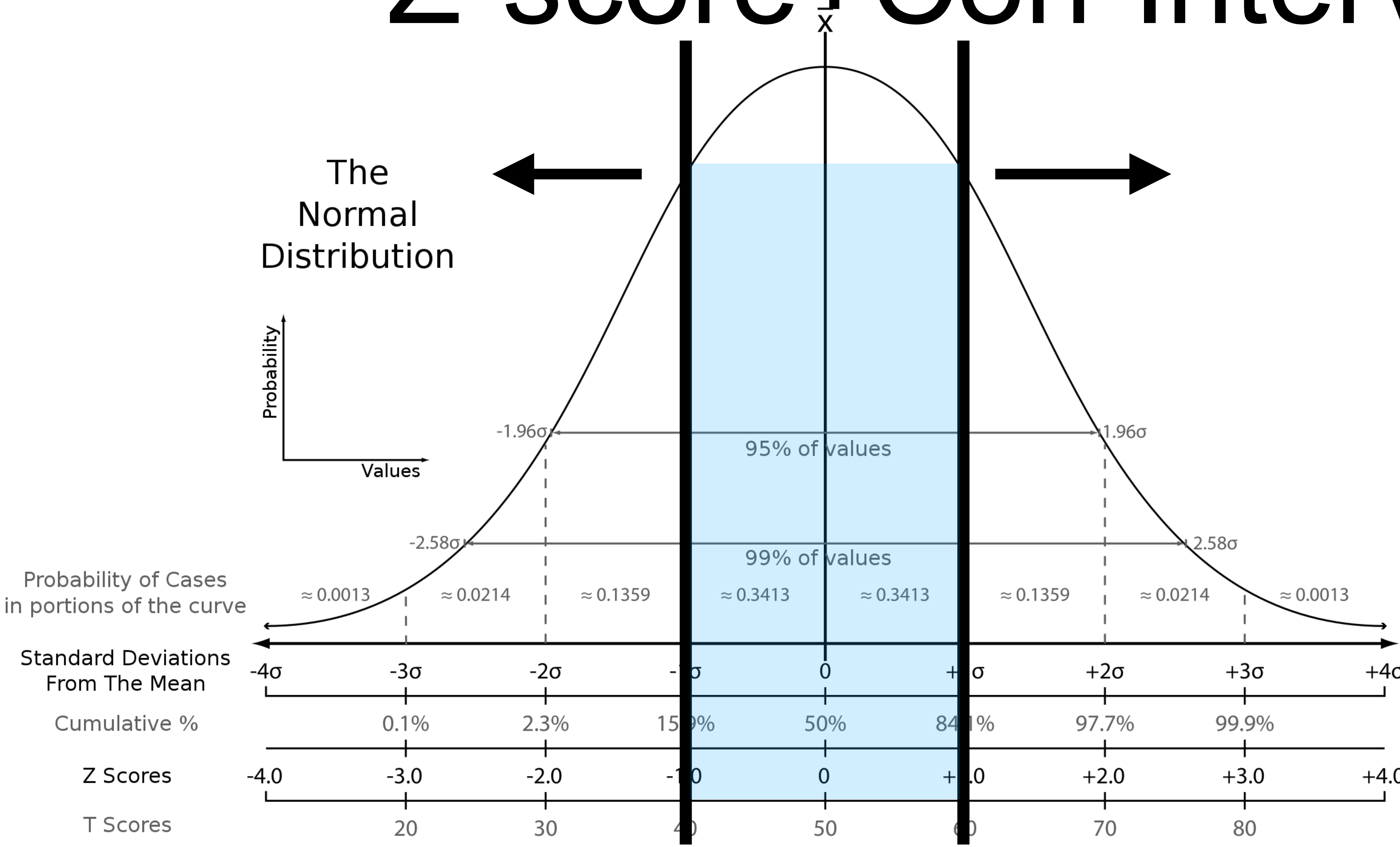


Confidence Intervals

Z-score+Con-Interval

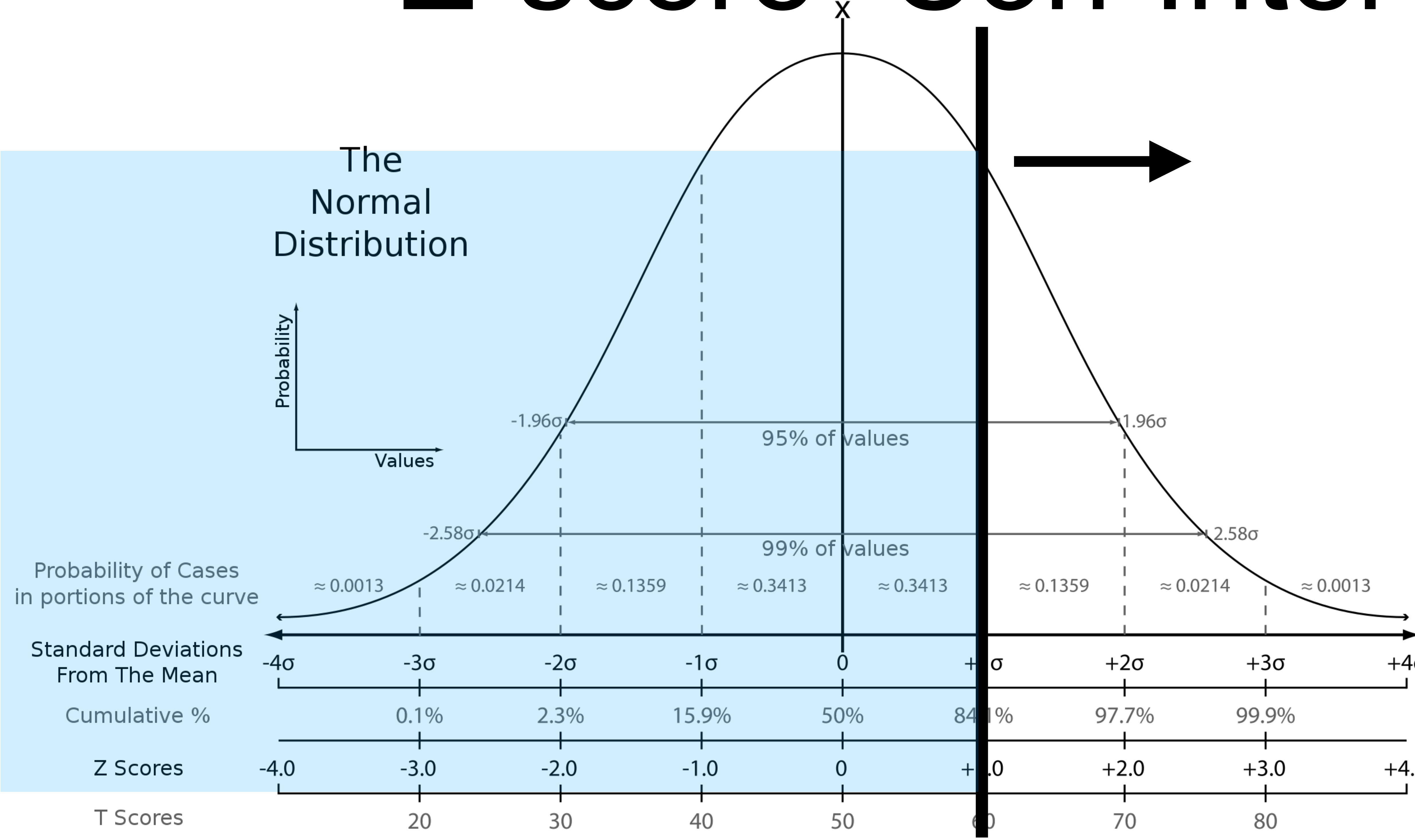


Z-score+Con-Interval



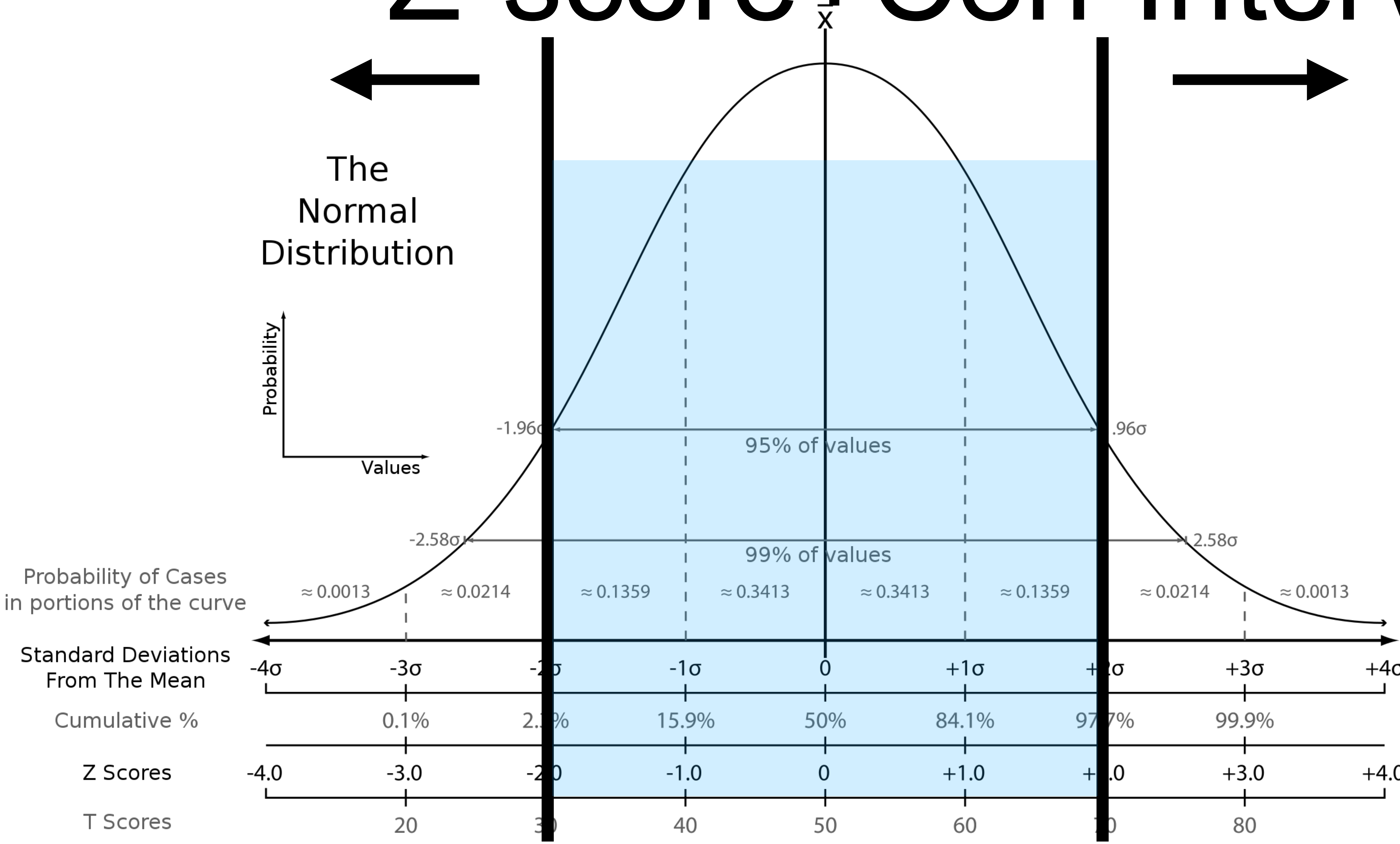
Z-score of 1: 68% chance of being within 1 standard deivation

Z-score+Con-Interval



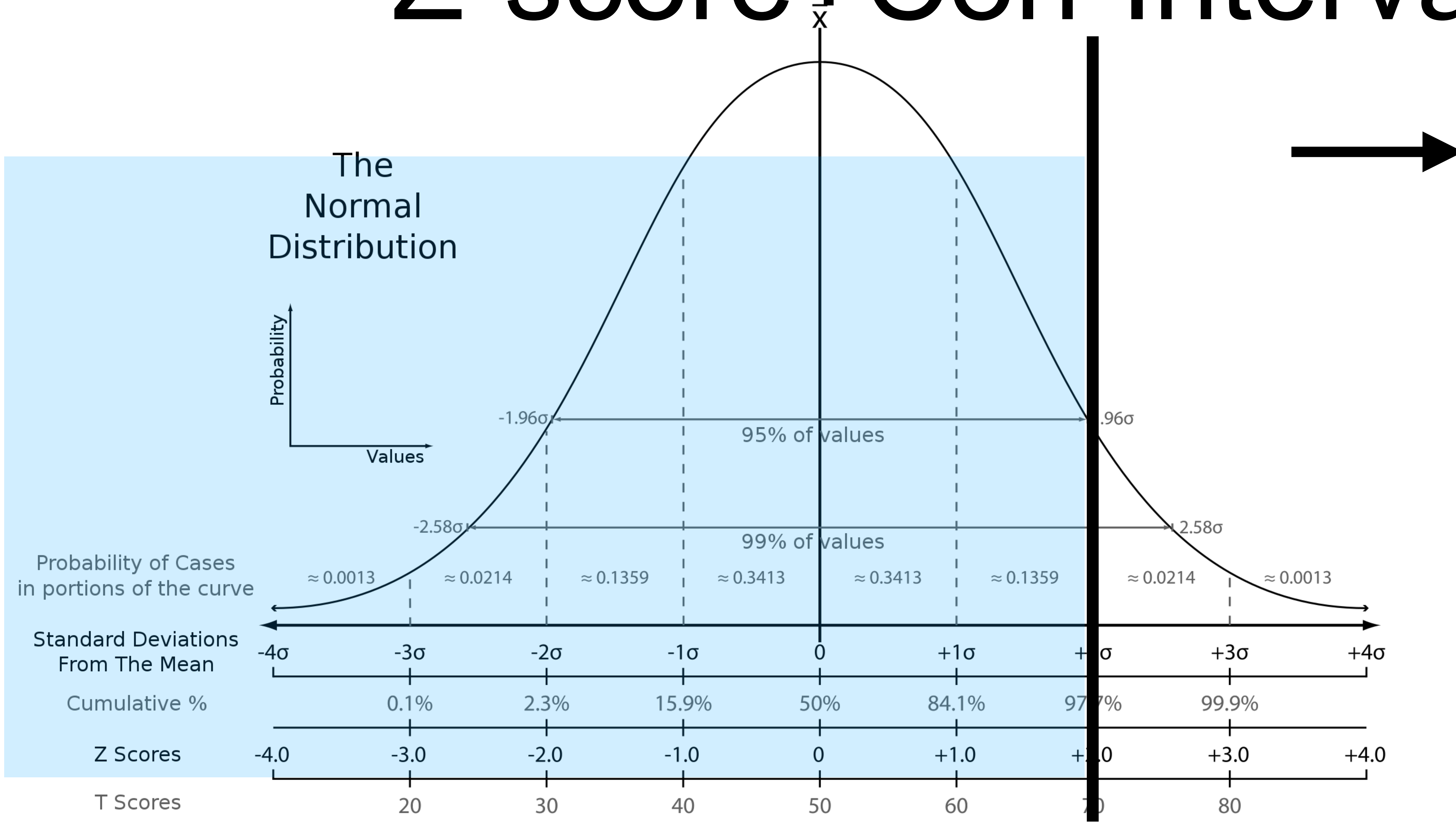
Single sided Z-score of 1: 84% chance of being above 1 standard deviation

Z-score+Con-Interval



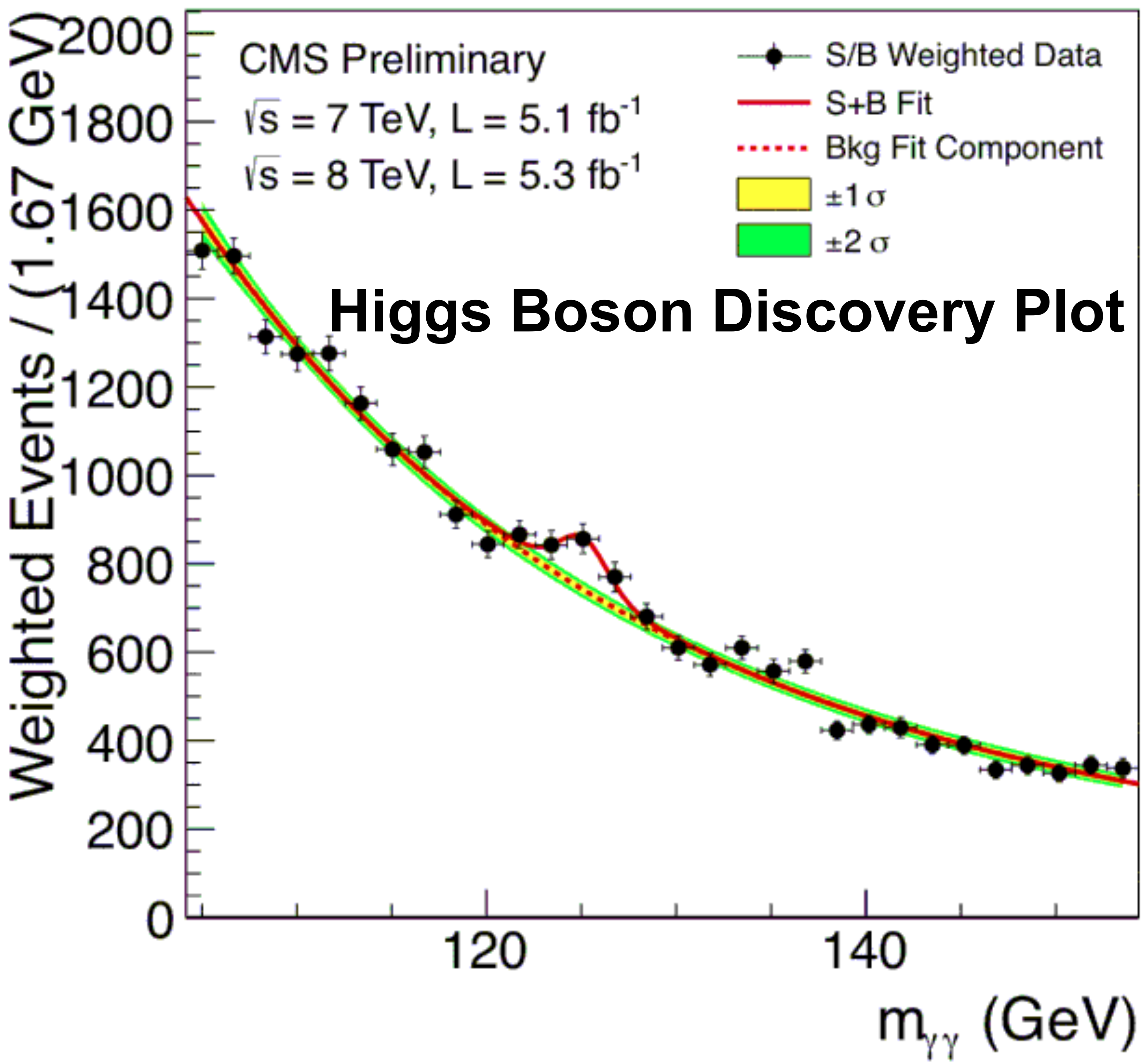
Z-score of 2: 95% chance of being within 2 standard deivation

Z-score+Con-Interval



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

Confidence Plots

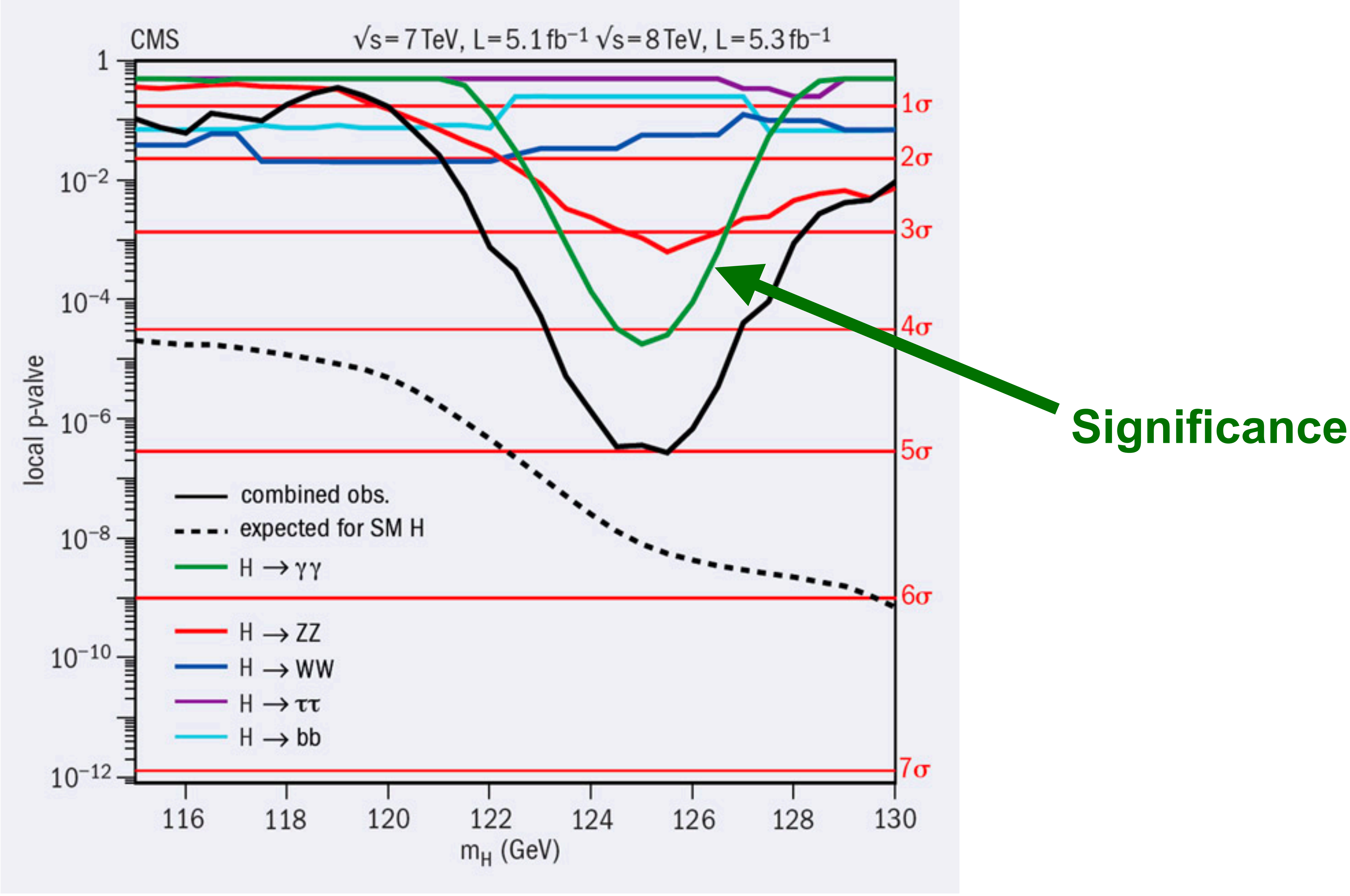


The rules of significance

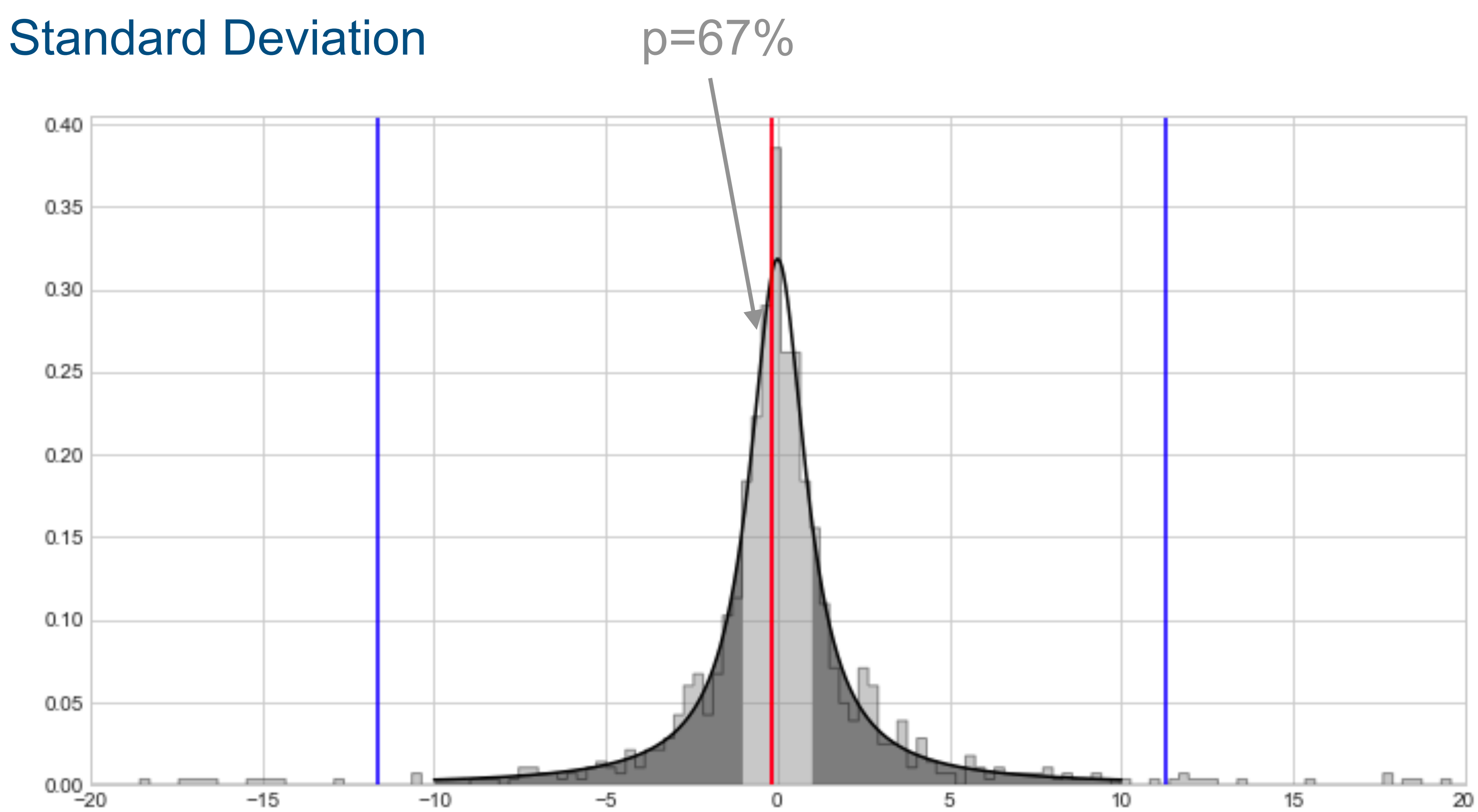
- How significant is our measurement? (High Energy physics rules)
- 3 sigma is considered “Evidence”
- 5 sigma is considered “Discovery”

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

Confidence Plots



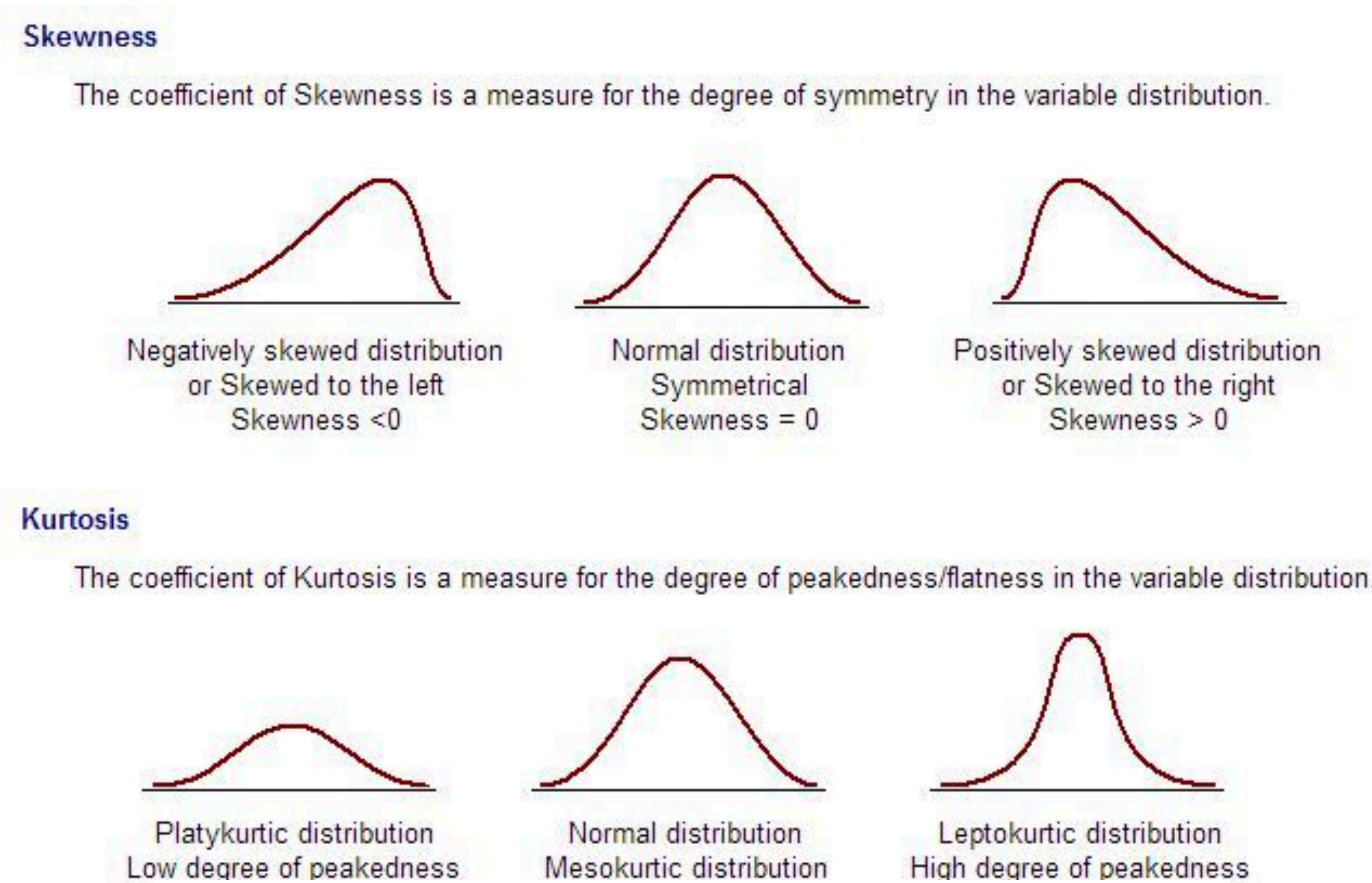
Z-score+Con-Interval



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

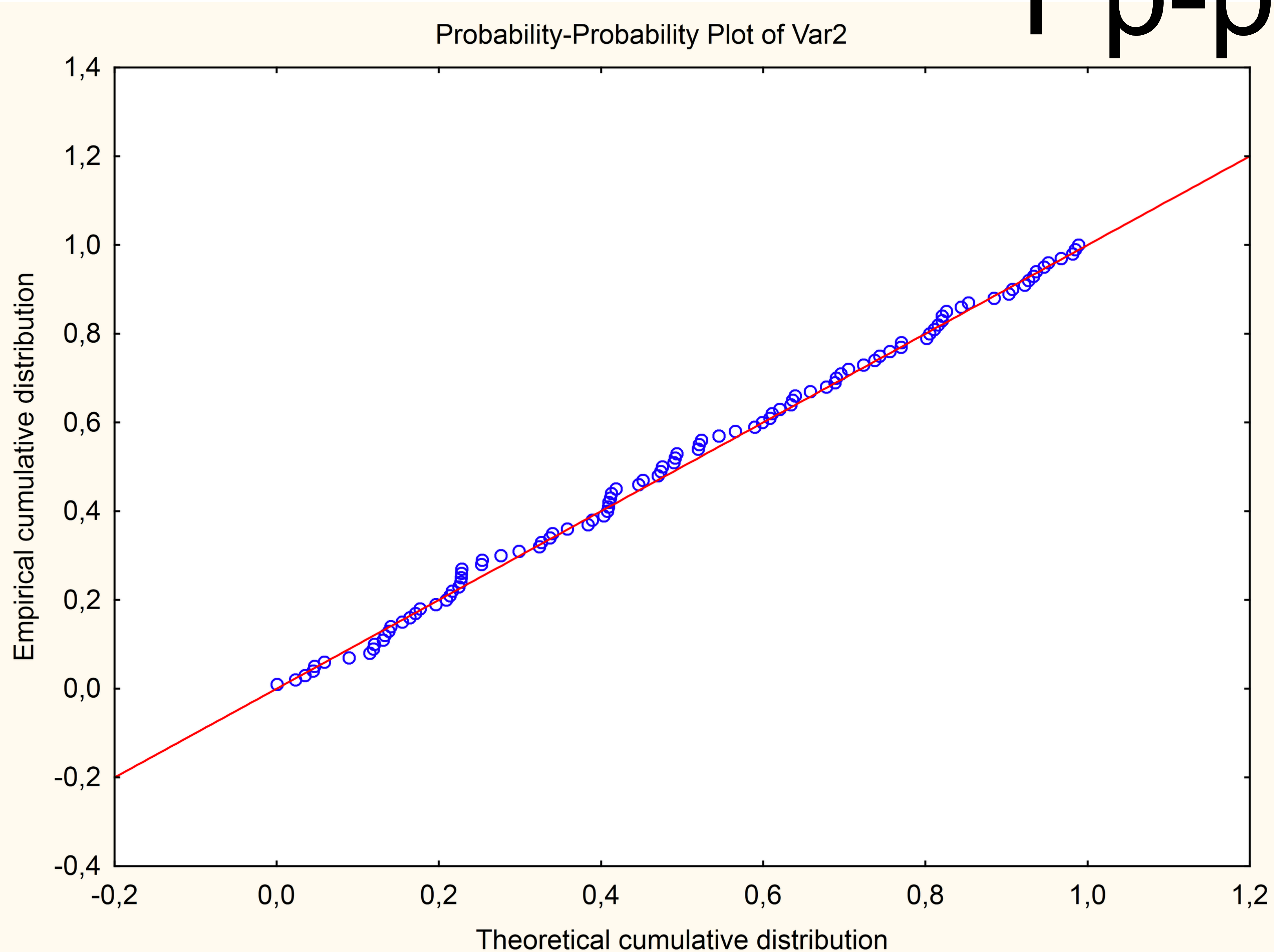
Moments

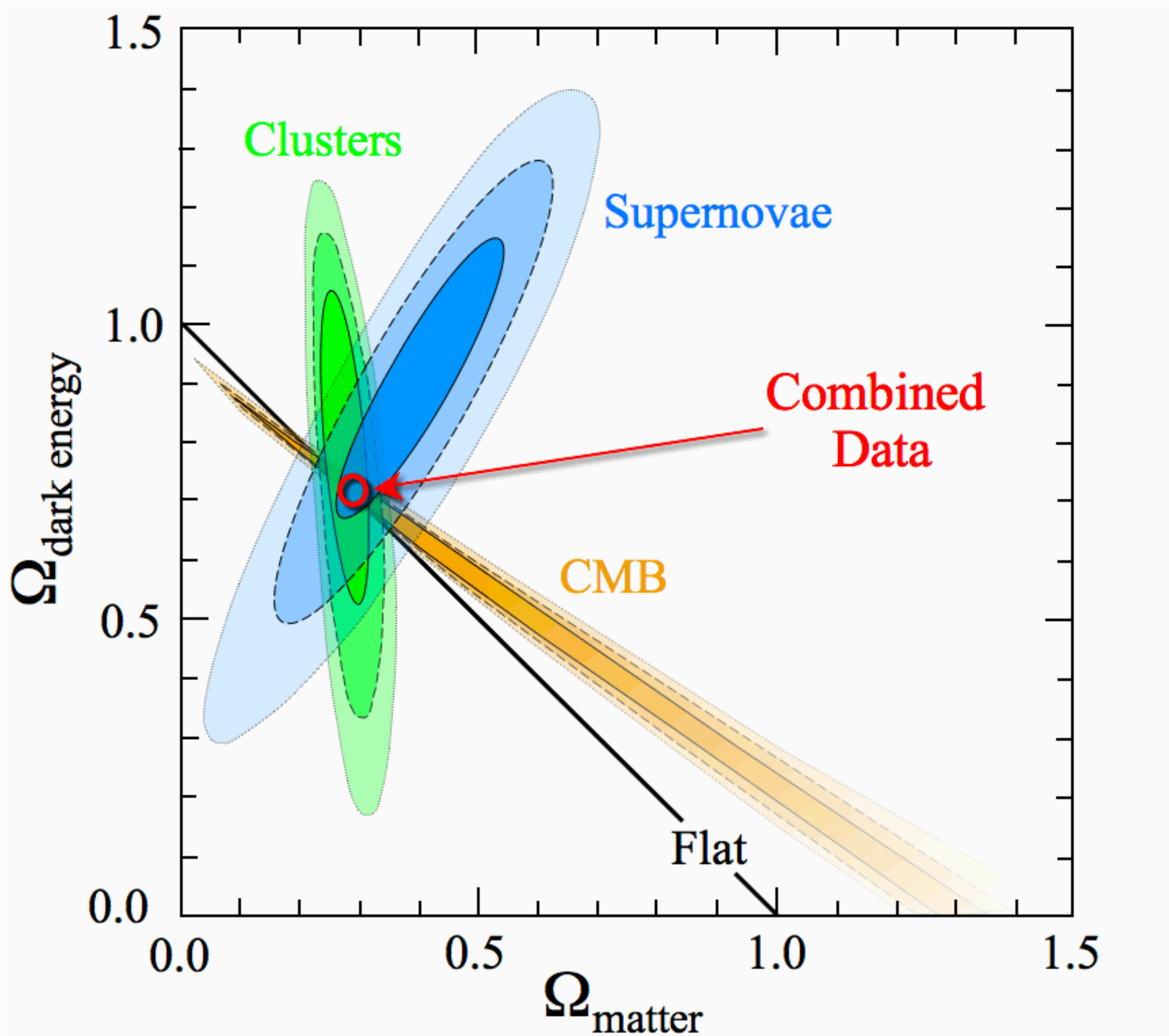
$$\mu_n = m^n(x) = E[x^n p(x)] = \int_{-\infty}^{\infty} x^n p(x) dx$$



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

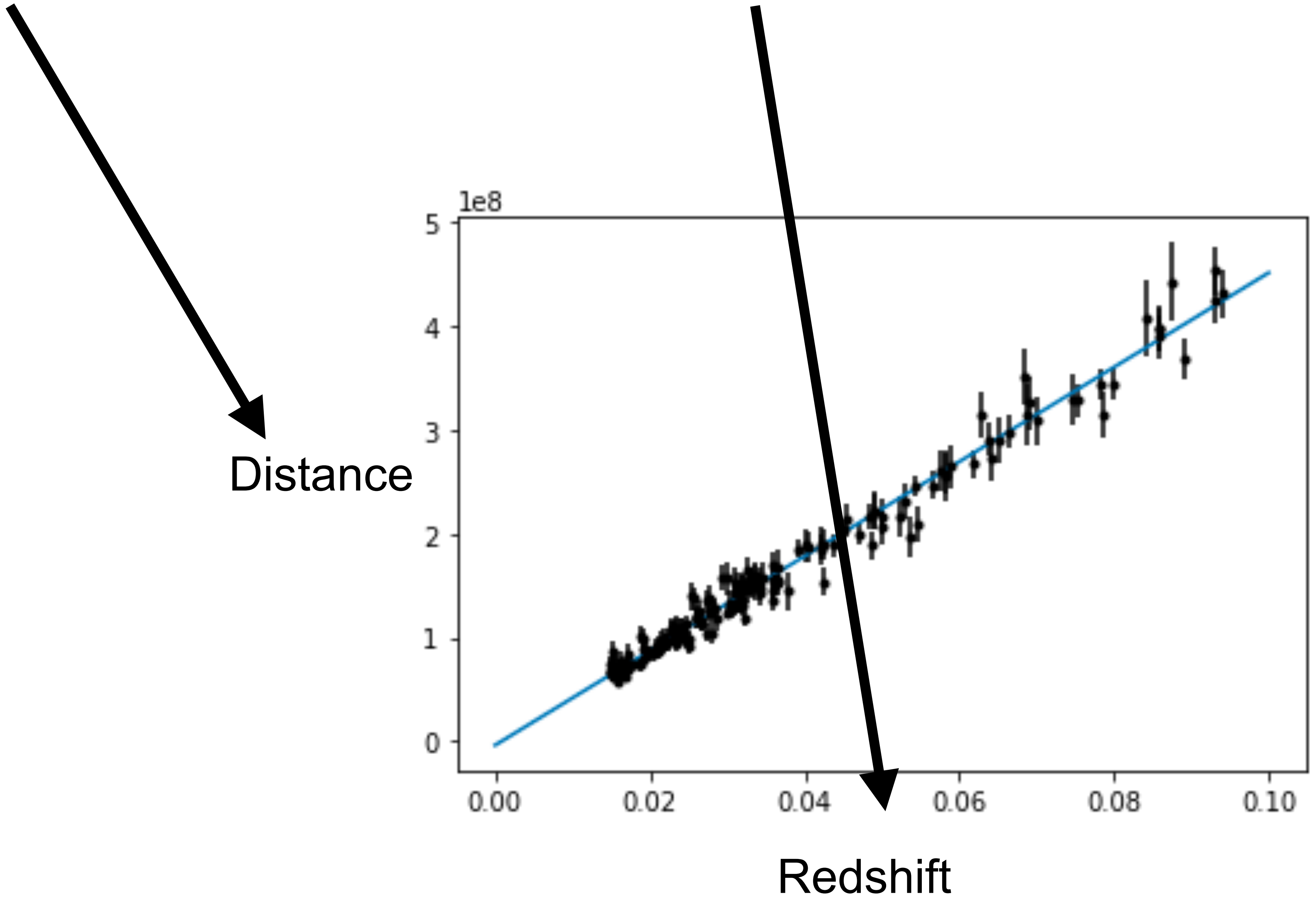
Pp-plot


Probability-Probability Plot of Var2
Empirical cumulative distribution
Theoretical cumulative distribution
1,4
1,2
1,0
0,8
0,6
0,4
0,2
0,0
-0,2
-0,4
-0,2
0,0
0,2
0,4
0,6
0,8
1,0
1,2



Clusters
Supernovae
Combined Data
CMB
Flat
1.5
1.0
0.5
0.0
0.0
0.5
1.0
1.5
$\Omega_{\text{dark energy}}$
Ω_{matter}

Friedmann Equations



Friedmann Equations

$$\left(\frac{h}{h_0}\right)^2 = (\Omega_m + \Omega_{\rm DM})a^{-3} + \Omega_r a^{-4} + \Omega_\kappa a^{-2} + \Omega_\Lambda$$

$$\boxed{1 = (\Omega_m + \Omega_{\rm DM}) + \Omega_r + \Omega_\kappa + \Omega_\Lambda}$$

Formula at time now

$$\boxed{1 = \Omega_M + \Omega_\Lambda \text{ or } \Omega_\Lambda = 1 - \Omega_M}$$

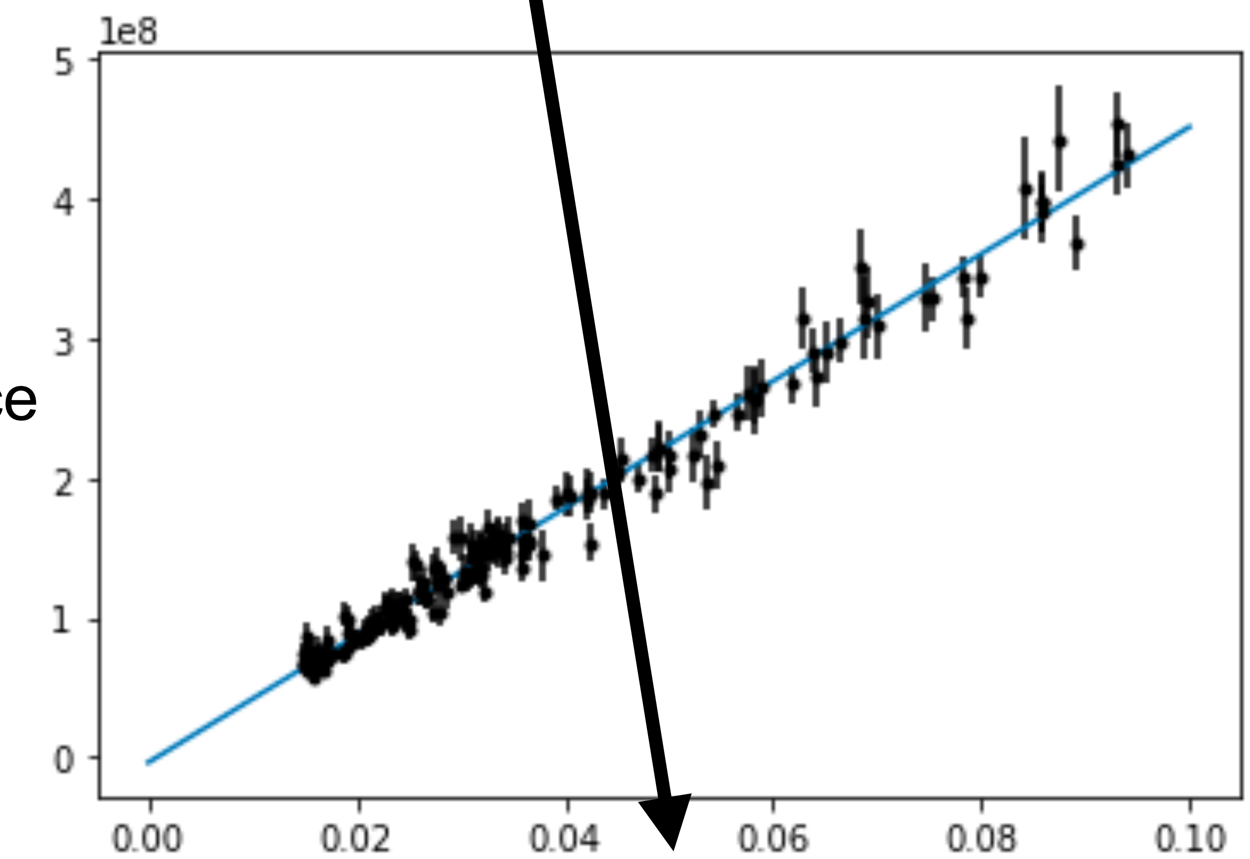
Removing Terms

$$\boxed{\left(\frac{h}{h_0}\right)^2 = (\Omega_M)a^{-3} + 1 - \Omega_M}$$

Friedmann Equations

$$d(z) = ct' = (1+z)ct = (1+z)\frac{c}{h_0}\int_0^z \frac{dz'}{\sqrt{\Omega_M(1+z')^3 + 1 - \Omega_M}}$$

Distance



Redshift

Our final expansion Plot