A Method for Upper Limit Calculation

Zhe Wang
(Thanks for the courtesy of Hanyu Wei)
May 11, 2017

- Last time we talked about TMinuit which is actually the engine of all root fit function.
- The manual is at http://seal.web.cern.ch/seal/documents/minuit/mnusersguide.pdf
- http://hep.fi.infn.it/minuit.pdf
- http://www-glast.slac.stanford.edu/software/root/GRUG/docs/Feature/GRUGminuit.pdf
- Originally it is in Fortran
- Root added a C++ wrapping for several applications
- The last reference will lead you to a quick start.

Outline

- Foreword
- The Method Basis
  - A test statistic
  - The distribution
  - upper limit
- Upper limit calculation procedure
- A simple study on the Method
- Real application

Foreword

- (what is it) a concise method for upper limit calculation based on asymptotic formula for likelihood-based test
- (why is it)
  - be free of computationally expensive MC calculation
  - Be able to report error bands of the upper limit
  - It is frequentist approach. No bias from a priori assumption with Bayesian approach
  - It can consider any shape of the signal. Avoid the ambiguity of selecting a signal region.
  - What's more, it is used widely in some discovery analysis, e.g., the Higgs searching
- (how to implement it)
  - Real data
  - Asimov data (expected data for background only)
  - Signal form we look for
  - Likelihood-based test statistic
  - ROOT

Test statistic (1)

Suppose one measures a physical quantity, and we use these values to construct a histogram $\mathbf{n}=(n_1, \ldots, n_N)$. The expectation value of n_i can be written

$$E[n_i] = \mu s_i + b_i$$

where b_i denotes the bkg events and s_i denotes the signal events in the i_{th} bin.

If s_i is normalized and contains all the signal shape information, we can take μ as a strength parameter. This is what we do generally and the strength parameter is important in the following test statistic.

Test statistic (2)

If the expected value of the histogram is also determined by some other parameters besides the strength parameter we mentioned before, we use $\boldsymbol{\theta}=(\boldsymbol{\theta}_1, \boldsymbol{\theta}_2, \ldots)$ to denote and we call them nuisance parameters.

Now to test a hypothesized μ we consider the profile likelihood ratio

$$\lambda(\mu) = \frac{L(\mu, \hat{\hat{\theta}})}{L(\hat{\mu}, \hat{\theta})}$$

Here there's no special requirement of the likelihood and $\hat{\hat{\theta}}$ in the numerator denotes the value of θ that maximizes L for the specified and fixed μ which we set in the likelihood under test. Note *θ is free.* While the denominator is the maximized (unconditional) likelihood function and all the parameters are free (the best fit).

Test statistic (3)

From the definition of $\lambda(\mu)$, one can see that $0 \le \lambda \le 1$, with near 1 implying good agreement between the data and the hypothesized value of μ *(we set and fix)*. Equivalently it is convenient to use the statistic

$$\boxed{q_{\mu} = -2\ln\lambda(\mu) = -2\ln\frac{L(\mu,\hat{\hat{\theta}})}{L(\hat{\mu},\hat{\theta})}}$$

We can see that higher values of q_{μ} thus correspond to increasing incompatibility between the data (best fit) and μ *(we set and fix)*.

This test statistic is what we will use for upper limit calculation and it has some great properties.

Approximate distribution (1)

Consider a test of the strength parameter μ and suppose the data are distributed according to a strength parameter μ' (true value). The desired distribution $f(q_\mu|\mu')$ can be found.

Firstly, the test statistic has an approximate form (asymptotic)

$$q_\mu = -2\ln\lambda(\mu) = -2\ln\frac{L(\mu,\hat{\hat{\theta}})}{L(\hat{\mu},\hat{\theta})} = \frac{(\mu-\hat{\mu})^2}{\sigma^2} + \mathrm{O}(1/\sqrt{N})$$

Here $\hat{\mu}$ is the best fit to data. *μ is what we set and fix there and the value we want to test and has nothing to do with* μ'. Generally $\mu \neq \mu'$

More importantly, $\hat{\mu}$ follows a Gaussian distribution with a mean μ' (true value) and standard deviation σ, and N represents the data sample size.

Approximate distribution (2)

From the approximate form above, we can find the distribution called non-central chi-square distribution for one d.o.f.

$$f(q_\mu \mid \mu') = \frac{1}{2\sqrt{q_\mu}} \frac{1}{\sqrt{2\pi}} \left[\exp\left(-\frac{1}{2}(\sqrt{q_\mu} + \sqrt{\Lambda})^2 \right) + \exp\left(-\frac{1}{2}(\sqrt{q_\mu} - \sqrt{\Lambda})^2 \right) \right]$$

Where the non-centrality parameter Λ is

$$\Lambda = \frac{(\mu - \mu')^2}{\sigma^2}$$

Here q_μ is the random variable as it contains a random variable $\hat{\mu}$, but we should pay attention that it is the value of q_μ has the so called non-central chi-square distribution.

μ is the test value we set by ourselves and μ' is the true value of strength.

For the special case $\mu' = \mu$ *one has* $\Lambda=0$ and $q_\mu = -2\ln\lambda(\mu)$ approaches a chi-square distribution for one degree of freedom.

Upper Limit (1)

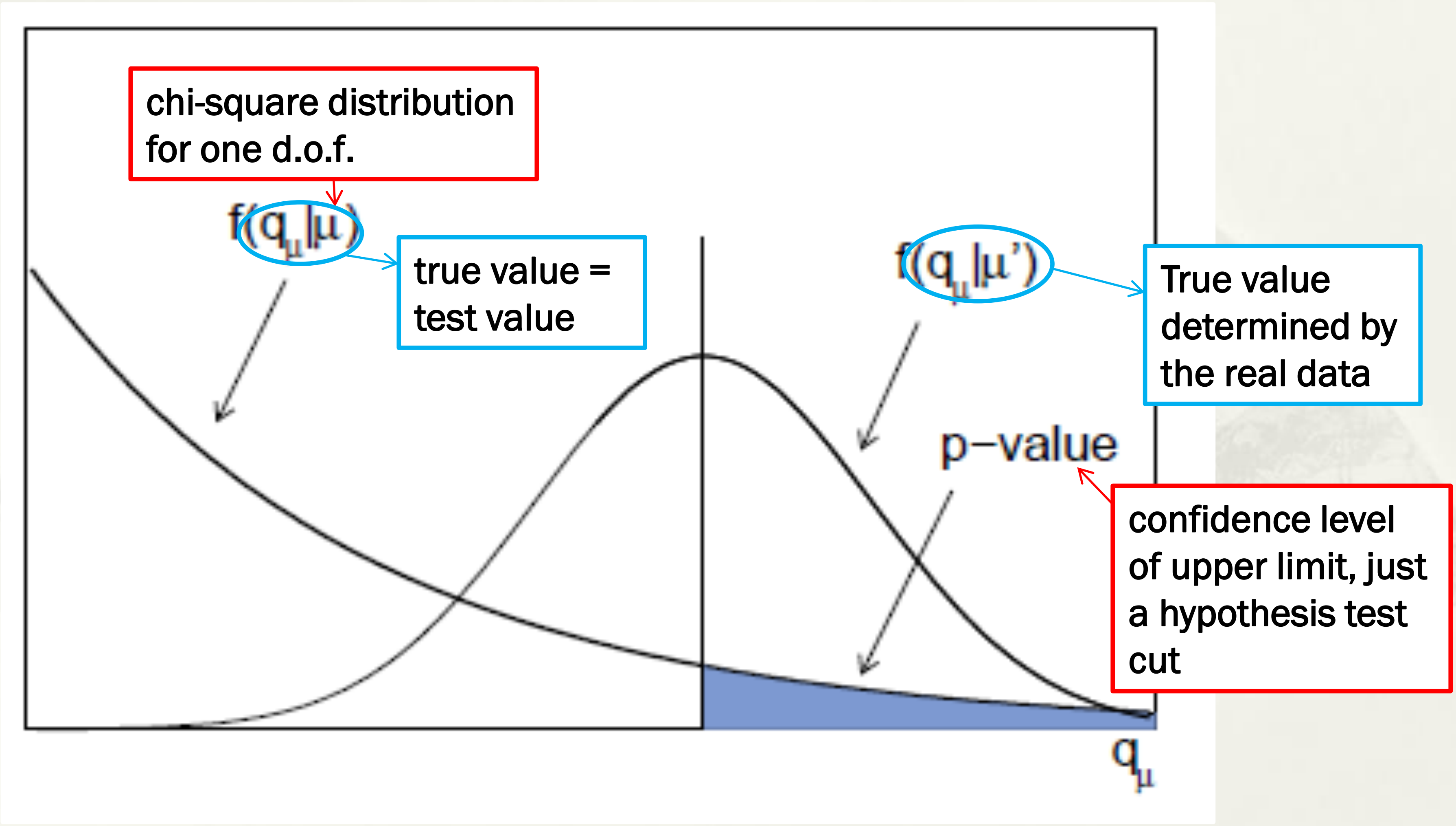
- *Upper limit* indicates the power of rejection (or reception). It'll give us the information about how likely there will be a signal and how significant to us.
- For a true value μ' of strength parameter, test statistic distribution $f(q_\mu|\mu')$ varies with different test value *μ* which is *we put in and fixed.* It has nothing to do with μ' .
- Change μ and compare distribution $f(q_\mu|\mu)$ with $f(q_\mu|\mu')$, we can get a certain *μ* meets some requirements and that is an upper limit we want.

Upper Limit (2)

item	formula	value	comment
Likelihood ratio	$\lambda(\mu)=\dfrac{L(\mu,\hat{\hat{\theta}})}{L(\hat{\mu},\hat{\theta})}$	0~1	0: bad agreement 1: good agreement
Test statistic	$q_\mu=-2\ln\lambda(\mu)$	0~∞	$q_\mu \approx \dfrac{(\mu-\hat{\mu})^2}{\sigma^2}$ random variable μ' hidden
Distribution	$f(q_\mu \mid \mu')=\dfrac{1}{2\sqrt{q_\mu}}\dfrac{1}{\sqrt{2\pi}}\left[\exp\left(-\dfrac{1}{2}(\sqrt{q_\mu}+\sqrt{\Lambda})^2\right)+\exp\left(-\dfrac{1}{2}(\sqrt{q_\mu}-\sqrt{\Lambda})^2\right)\right]$		

$$\boxed{\Lambda=\frac{(\mu-\mu')^2}{\sigma^2}}$$

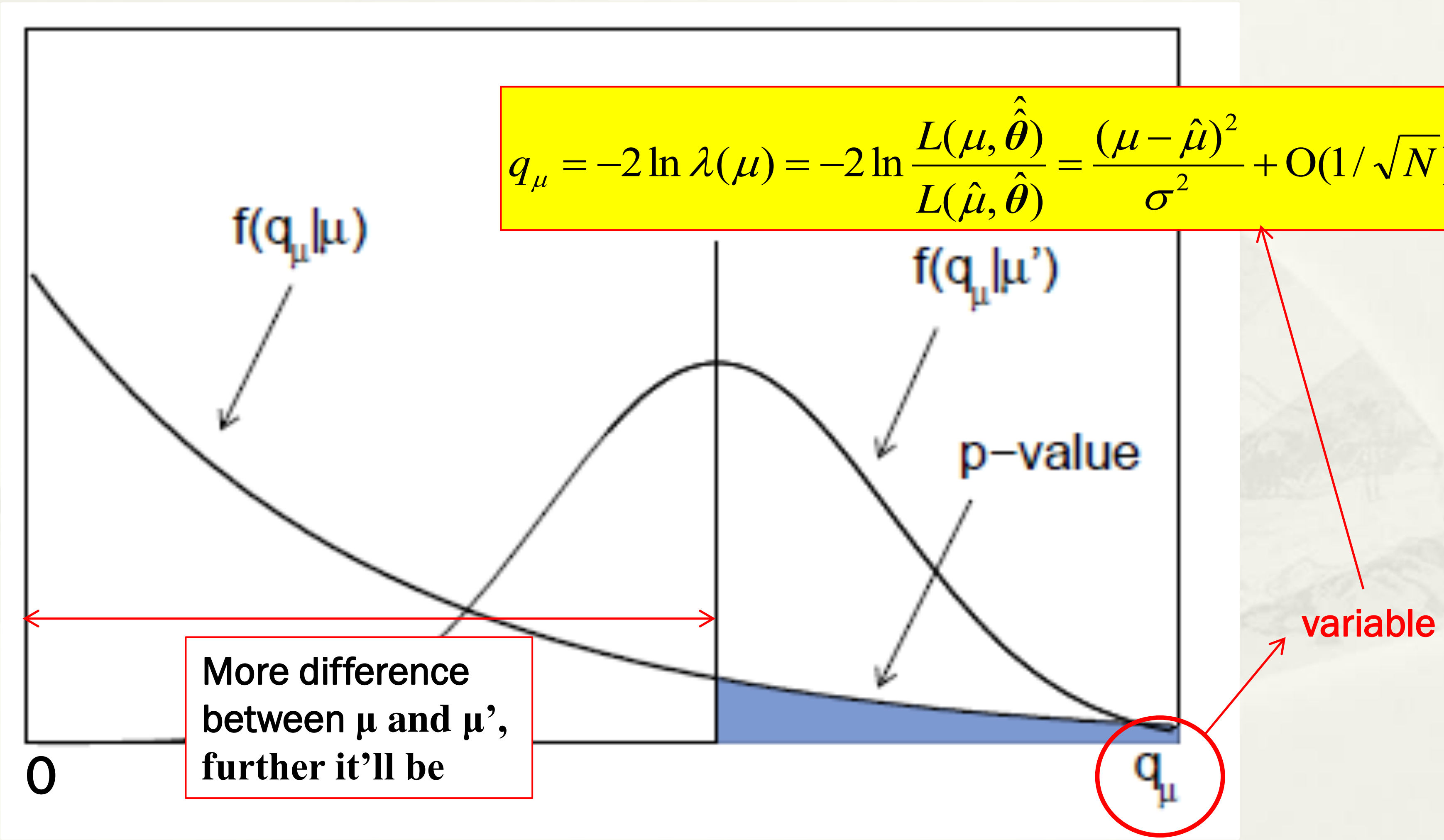
Upper Limit



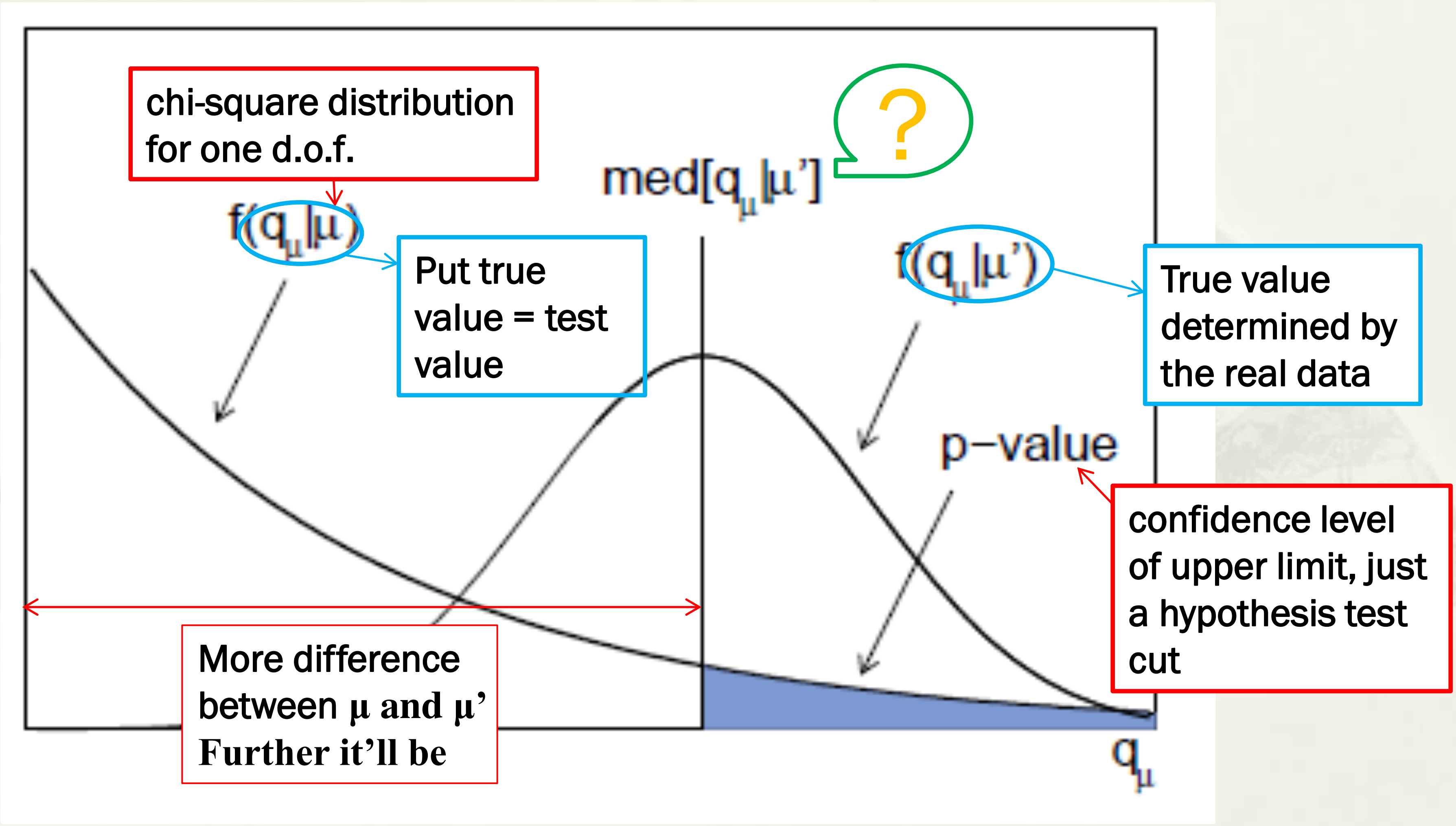
chi-square distribution for one d.o.f.
f(qμ|μ)
true value = test value
f(qμ|μ')
True value determined by the real data
p–value
confidence level of upper limit, just a hypothesis test cut
qμ

Upper Limit

$$q_\mu = -2\ln\lambda(\mu) = -2\ln\frac{L(\mu,\hat{\hat{\theta}})}{L(\hat{\mu},\hat{\theta})} = \frac{(\mu-\hat{\mu})^2}{\sigma^2} + O(1/\sqrt{N})$$



Upper Limit


chi-square distribution for one d.o.f.
f(q_μ|μ)
Put true value = test value
med[q_μ|μ']
?
f(q_μ|μ')
True value determined by the real data
p−value
confidence level of upper limit, just a hypothesis test cut
More difference between μ and μ' Further it'll be
q_μ

Upper Limit (2)

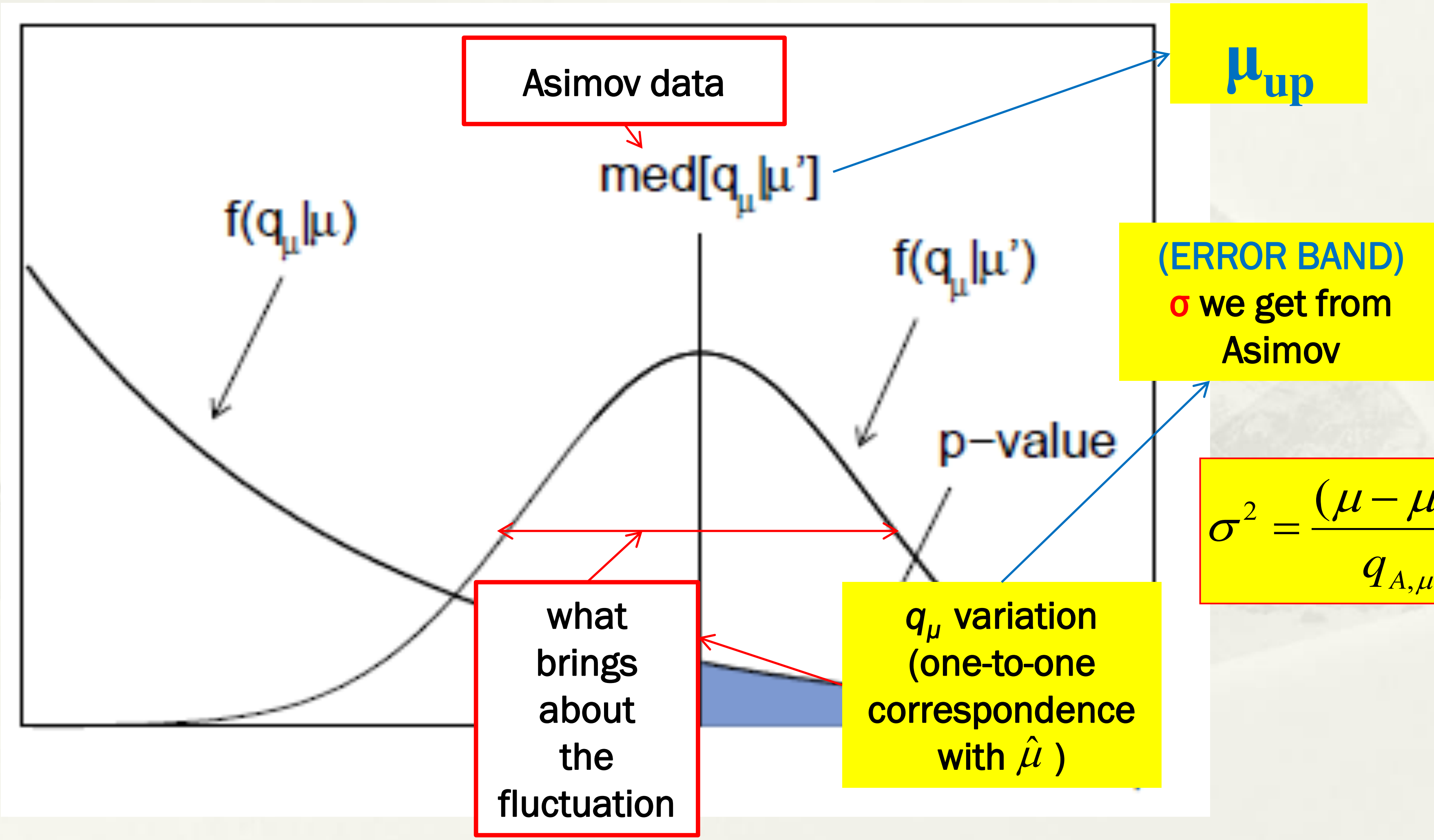
- Asimov data set
  - One uses it to evaluate the estimators for all parameters, one obtains the true parameter values. Generally Asimov data are equal to the expectation values. Median (i.e. mean, expected) value here corresponds to the Asimov data set

$$\lambda_A(\mu) = \frac{L_A(\mu, \hat{\hat{\theta}})}{L_A(\hat{\mu}, \hat{\theta})} = \frac{L_A(\mu, \hat{\hat{\theta}})}{L_A(\mu', \hat{\theta})}$$

- Use asymptotic formula of $\lambda(\mu)$, we can get σ which is the deviation of $\hat{\mu}$

$$q_{A,\mu} = -2\ln\lambda_A(\mu) \approx \frac{(\mu-\mu')^2}{\sigma^2} \longrightarrow \sigma^2 = \frac{(\mu-\mu')^2}{q_{A,\mu}}$$

Upper Limit



$$\sigma^2 = \frac{(\mu - \mu')^2}{q_{A,\mu}}$$

Upper Limit (3)

- Through a detailed proof and calculation, we get that the median upper limit with $\pm N\sigma$ error band is given by

$$\boxed{band_{N\sigma} = med[\mu_{up} \mid \mu'] \pm N\sigma}$$

- Stress that
  - For expected upper limit and error bands, use Asimov Data
  - Median upper Limit is obtained from the μ scan which leads to different q_μ (using quantile of chi-square distribution for one d.o.f.)
  - Error band corresponds to the fluctuation of $\hat{\mu}$ and we get it using the asymptotic relation between q_μ and $\hat{\mu}$
  - If you just want to get some real data's upper limit (observed upper limit), do a μ scan and make the chi-square variation from the best fit value meet some requirements, e.g. 90%C.L. demands delta chi^2 = 2.71. Here it is like just taking real data as Asimov data.

Upper Limit Calculation Procedure

- 1) Know the Asimov data, i.e. the expected value . If we are looking for a signal, we should know the background. However, if the signal is small enough we can use the histogram fit result as Asimov data (expected bkg value) expediently.
- 2) Construct a likelihood and then we have the test statistic
- 3) You set and fix hypothesized value of strength parameter μ and get the test statistic value for each μ until the value of test statistic meets the requirements (chi^2 quantile)
- 4) When we get the median μ_{up} from Step 3, use asymptotic formula to calculate the deviation of $\hat{\mu}$ at this time and that is the error band we want.
- 5) For the real data, use them instead of Asimov data and the same likelihood and method, then we'll get the observed upper limit.
- 6) Compare the observed value with expected value, it'll bring you some special information and direction

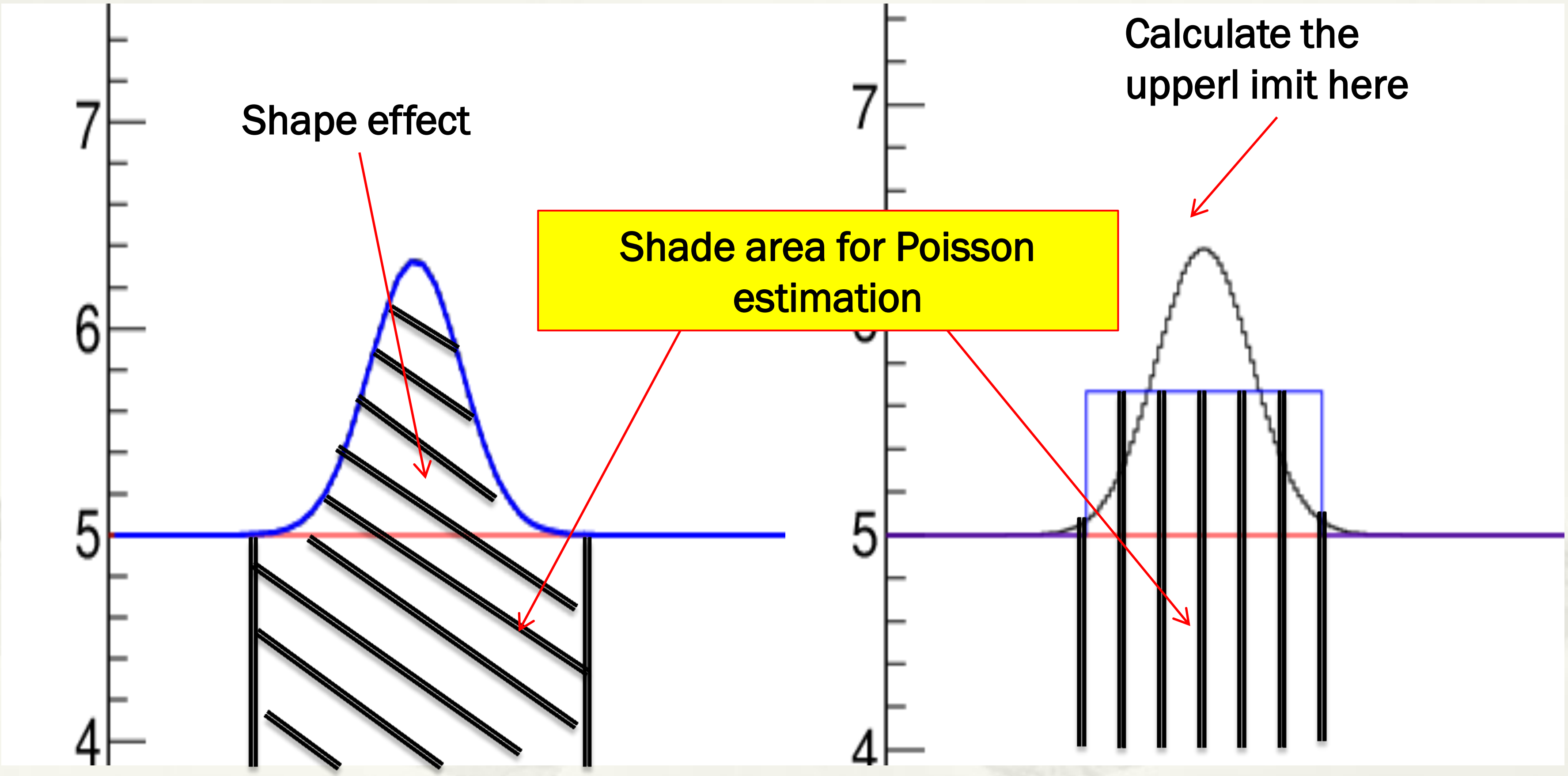
A Simple Study on the Method (1)

- We put a uniform background there and a Gaussian Signal at some place to test the method.
- In the likelihood, two kinds of test signal form (i.e. the expected signal form)
  - Gaussian with the same sigma as the real data we put
  - A uniform in some range, here we let the range be 5 sigma
- Likelihood (use the same likelihood in the following calculation)

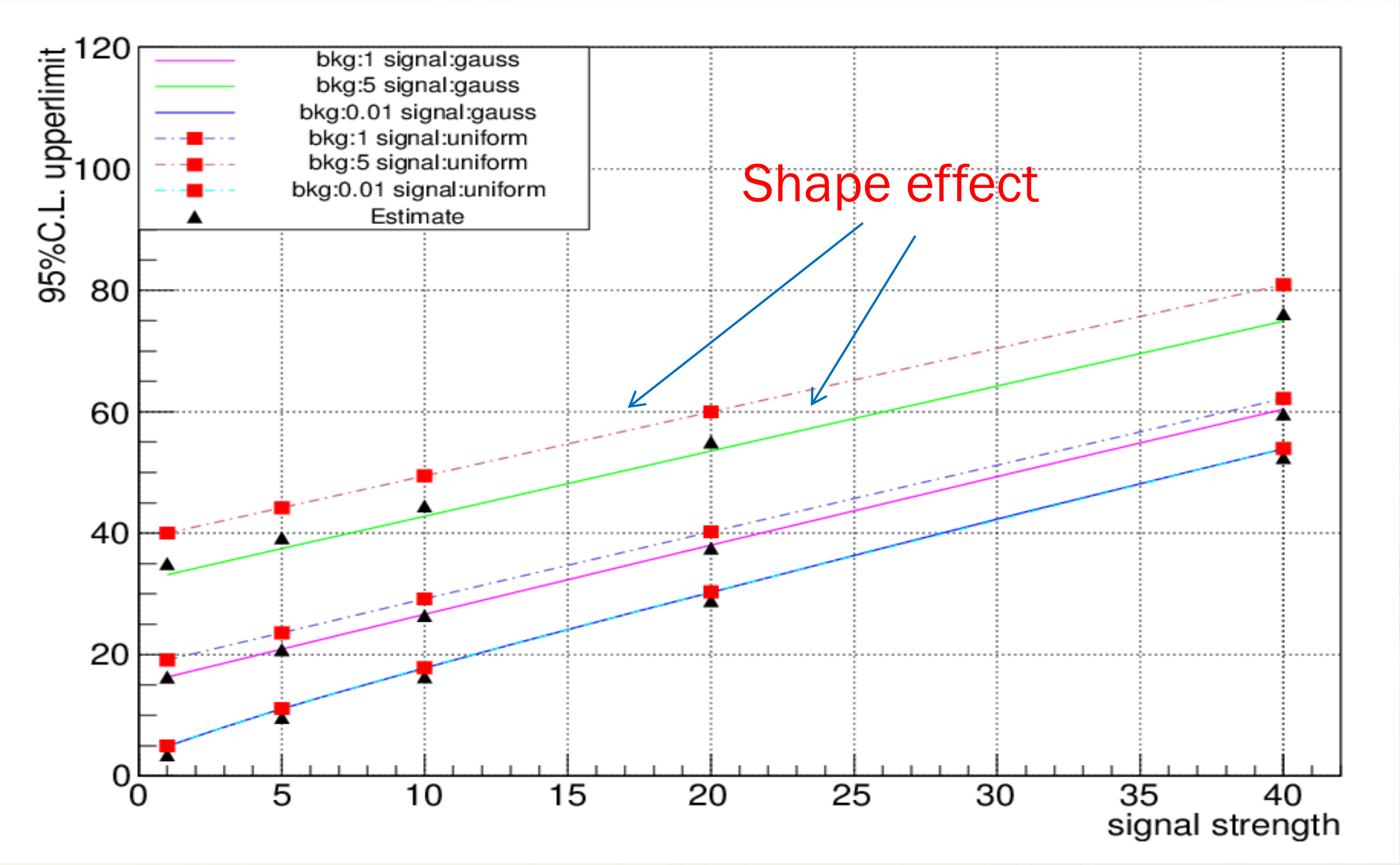
$$L(\mu,\theta)=\prod_{i}^{Nbins}\frac{\left(\mu\cdot s_i+\theta\cdot b_i\right)^n}{n!}e^{-\left(\mu\cdot s_i+\theta\cdot b_i\right)}$$

- We want to know how different the upper limit (at signal central) we get using this method from a simple estimation
  - Using Poisson estimation
  - Sqrt [Bkg (in 5 sigma, shape) + signal strength (integral value, the factor multiplied with a normalized Gaussian or other forms)]

A Simple Study on the Method (2)



A Simple Study on the Method (3)



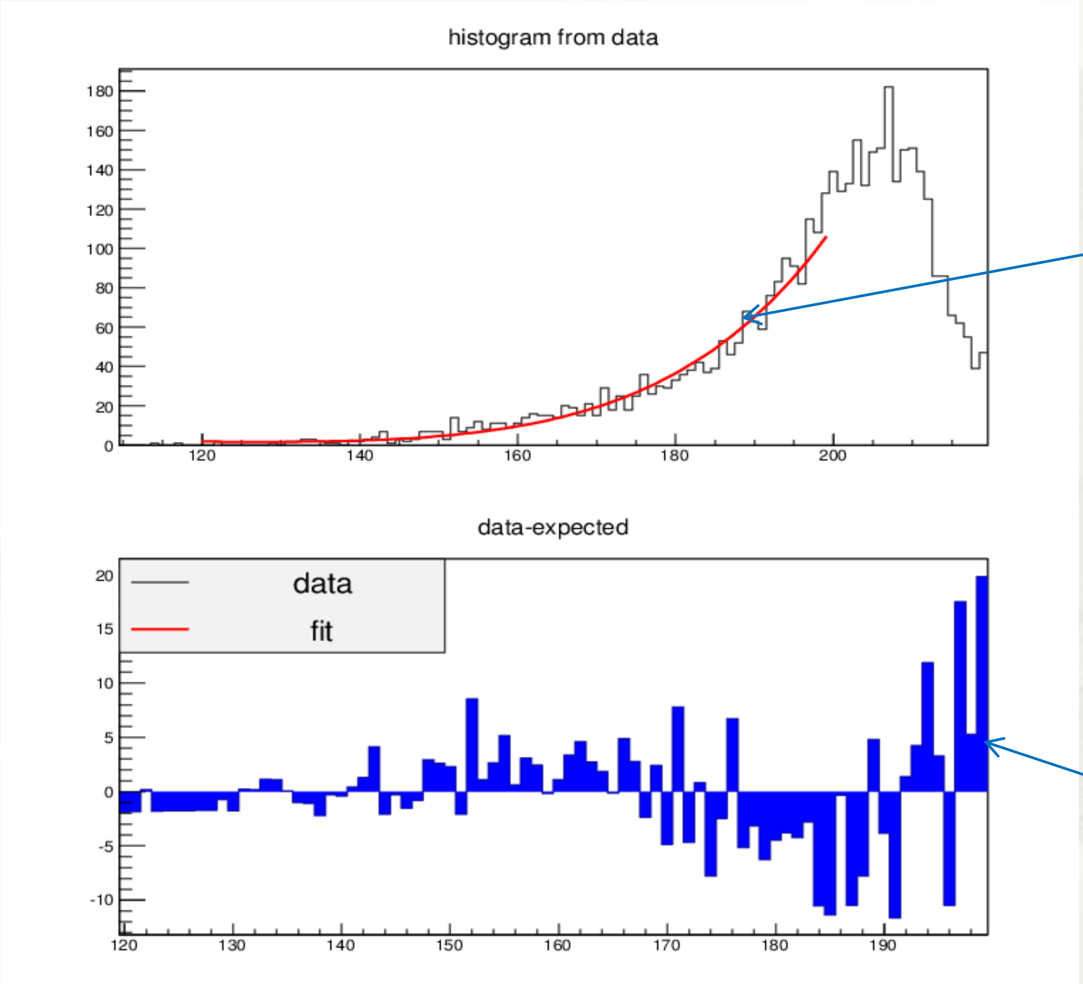
Application

- Heavy neutrino search in K decay (small signal search)
- Use the method in this talk to do analysis in every point in the Momentum spectrum
- The signal form we are looking for is a Gaussian with sigma ~3 MeV
- Likelihood is same as what we use above

$$L(\mu,\theta)=\prod_{i}^{Nbins}\frac{\left(\mu\cdot s_i+\theta\cdot b_i\right)^n}{n!}e^{-\left(\mu\cdot s_i+\theta\cdot b_i\right)}$$

- For simplicity here, we use histogram fit as Asimov data (background expected value) to calculate the expected upper limit value and error bands
- Use real data to calculate the observed upper limit
- Compare observed curve with the expected one to find where there may be a signal

Application



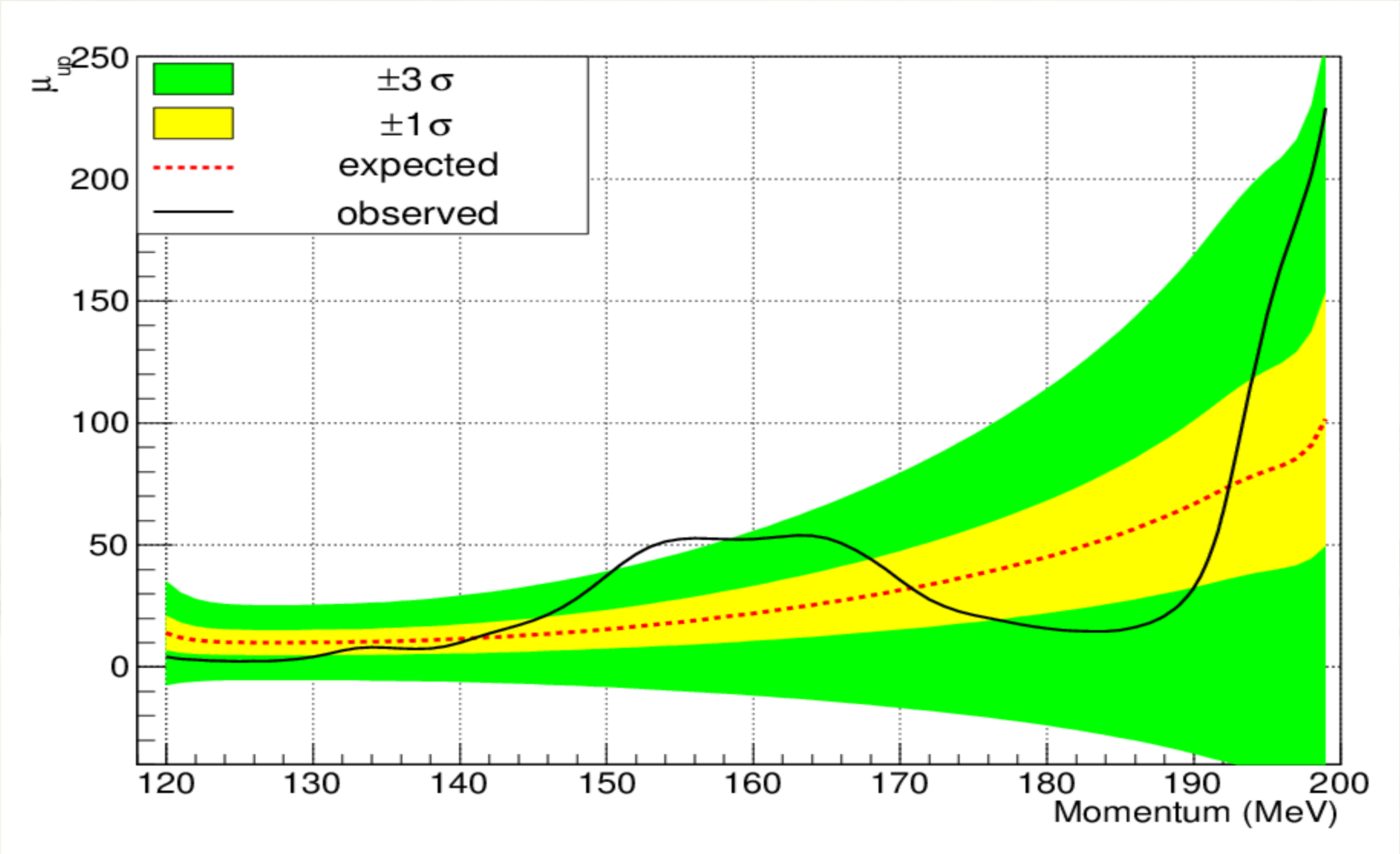
Momentum Spectrum Fit (120~200MeV)

Just use exp+pol5

Data histogram minus the Fit value

It should have the same tendency of observed upper limit

Application



Reference

Original Paper:

Asymptotic formulae for likelihood-based tests of new physics

arXiv:1007.1727v2 [physics, data-an] 3 Oct 2010

Glen Cowen, Kyle Cranmer, Eilam Gross, Ofer Vitells

Other material:

Statistical methods in experimental physics,

Sec. 9.4

Frederick James