Limit determination - Emphasis on arbitrariness

E. Busato

December 3, 2013

Outline

- Not addressed in this presentation
  - Shapes
  - Profiling

Standard cut&count problem

Notations

- One or more backgrounds: $\sum b_i$
- Signal with unknown x-sec: s (free parameter)
- Total expectation: $s + \sum b_i$
- Observed number of events: N_{obs}

- When N_{obs} agrees with background yield, infer upper limit s_{up}

$$\boxed{s_{\mathrm{up}} = \text{yield above which total expectation and observation disagree}}$$

- To set limit one needs to
  - Know how N_{obs} is distributed
  - Have a quantitative measure of the agreement between data and total expectation

Standard cut&count problem

Notations

- One or more backgrounds: $\sum b_i$
- Signal with unknown x-sec: s (free parameter)
- Total expectation: $s + \sum b_i$
- Observed number of events: N_{obs}

- When N_{obs} agrees with background yield, infer upper limit s_{up}

$$\boxed{s_{\text{up}} = \text{yield above which total expectation and observation disagree}}$$

- To set limit one needs to
  - Know how N_{obs} is distributed
  - Have a quantitative measure of the agreement between data and total expectation

Standard cut&count problem

Notations

- One or more backgrounds: $\sum b_i$
- Signal with unknown x-sec: s (free parameter)
- Total expectation: $s + \sum b_i$
- Observed number of events: N_{obs}

- When N_{obs} agrees with background yield, infer upper limit s_{up}

$$\boxed{s_{\text{up}} = \text{yield above which total expectation and observation disagree}}$$

- To set limit one needs to
  - Know how N_{obs} is distributed
  - Have a quantitative measure of the agreement between data and total expectation

Distribution of N_{obs} - Terminology

- Distribution of observation described by likelihood $\mathscr{L}$:

$$\mathscr{L} = \text{proba}(N_{\text{obs}}; s + \sum_i b_i)$$

- Likelihood also called statistical model or model

Two central problems to be addressed

1. Build statistical model
2. Infer s_{up} from statistical model (i.e. quantify agreement between observation and total expectation for each s)

Important point: no unique solution to these problems

$\rightarrow$ upper limit is a somewhat subjective notion (always relative to the choices made to solve these problems)

Two central problems to be addressed

1. Build statistical model
2. Infer s_{up} from statistical model (i.e. quantify agreement between observation and total expectation for each s)

Important point: no unique solution to these problems

$\rightarrow$ upper limit is a somewhat subjective notion (always relative to the choices made to solve these problems)

Softwares

- Throughout the presentation, three softwares implementing solutions to previous problems will be discussed/compared/studied
  - McLimit: hybrid frequentist-bayesian
  - CLsGenerator: hybrid frequentist-bayesian
  - BayesianMCMC: bayesian (Markov Chain Monte Carlo)
- HistFactory and various RooStats implementations (almost) not described

CLsGenerator and BayesianMCMC

- Use exactly the same statistical model
  - Difference is on how inference is made
- Take same files as input
  - Comparison straightforward

Single channel - No uncertainties

Statistical model and inference

1. **Statistical model**

   In our experiments, $N_{\text{obs}} \sim \textit{Binomial}$[1] $\rightarrow$ Approximated by Poisson

$$\boxed{\mathscr{L} = \frac{\left(s + \sum_i b_i\right)^{N_{\text{obs}}}}{N_{\text{obs}}!} e^{-\left(s + \sum_i b_i\right)}}$$

2. **Inference of s_{up}**

   - Two families of methods: frequentist and bayesian

[1] $\sim$ means "is distributed according to"

Statistical model and inference

1. **Statistical model**

   In our experiments, $N_{\text{obs}} \sim \textit{Binomial}$[1] $\rightarrow$ Approximated by Poisson

$$\boxed{\mathscr{L} = \frac{\left(s + \sum_i b_i\right)^{N_{\text{obs}}}}{N_{\text{obs}}!} e^{-\left(s + \sum_i b_i\right)}}$$

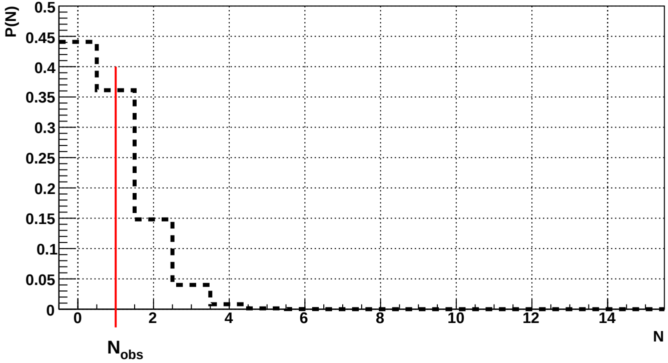
2. **Inference of s_{up}**

   - Two families of methods: frequentist and bayesian

[1] $\sim$ means "is distributed according to"

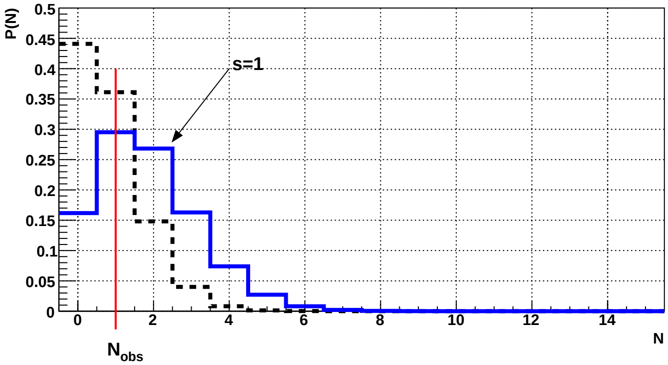
Frequentist solution

- Consider the case: $N_{\mathrm{obs}} = 1$ and $\sum_i b_i = 0.82$



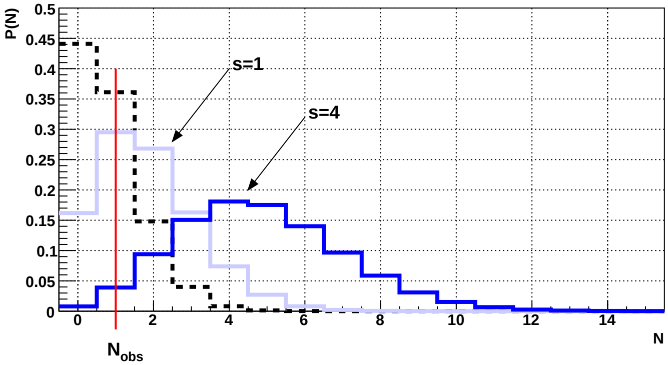
Frequentist solution

- Consider the case: $N_{obs} = 1$ and $\sum_i b_i = 0.82$



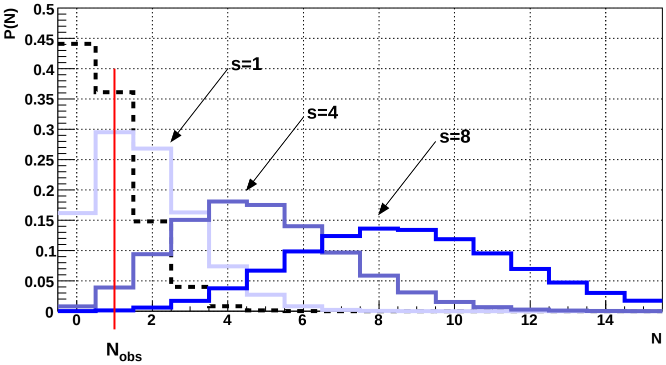
Frequentist solution

- Consider the case: $N_{\text{obs}} = 1$ and $\sum_i b_i = 0.82$



Frequentist solution

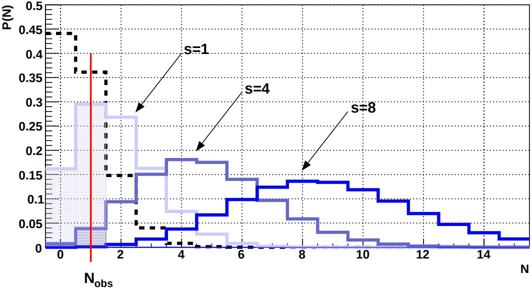
- Consider the case: $N_{\text{obs}} = 1$ and $\sum_i b_i = 0.82$



Frequentist solution - CL_{s+b} method

- Quantitative measure of agreement: $p-value$

$$p-value = \underbrace{\sum_{N=0}^{N_{obs}} \mathscr{P}(N; s+\sum_i b_i)}_{\text{c. d. f.}} = \sum_{N=0}^{N_{obs}} \frac{\left(s+\sum_i b_i\right)^N}{N!} e^{-\left(s+\sum_i b_i\right)}$$



- Remark: this $p-value$ is also called CL_{s+b}

Frequentist solution - CL_{s+b} method

- For which values of s does the observation disagree with total expectation ?
  - Usually we take values for which $p - value \leq \alpha$ with $\alpha = 0.05$
- Upper limit s_{up} is the smallest of all values for which observation and total expectation disagree

$$\boxed{CL_{s+b}(s_{\text{up}}) = \alpha}$$

- Previous example: $s_{\text{up}} = 3.92$

Frequentist solution - CL_{s+b} method

- For which values of s does the observation disagree with total expectation ?
  - Usually we take values for which $p - value \leq \alpha$ with $\alpha = 0.05$
- Upper limit s_{up} is the smallest of all values for which observation and total expectation disagree

$$\boxed{CL_{s+b}(s_{\text{up}}) = \alpha}$$

- Previous example: $s_{\text{up}} = 3.92$

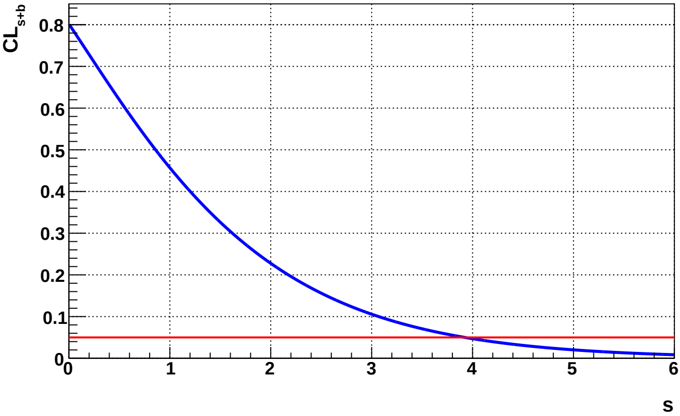
Frequentist solution - CL_{s+b} method

- For which values of s does the observation disagree with total expectation ?
  - Usually we take values for which $p - value \leq \alpha$ with $\alpha = 0.05$
- Upper limit s_{up} is the smallest of all values for which observation and total expectation disagree

$$\boxed{CL_{s+b}(s_{\mathrm{up}}) = \alpha}$$

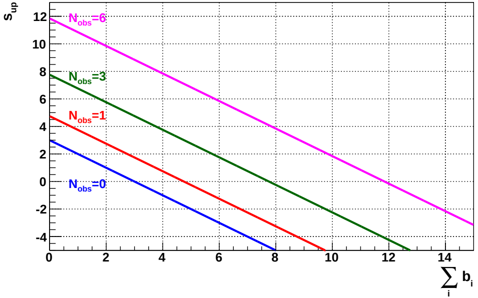
- Previous example: $s_{\mathrm{up}} = 3.92$

Frequentist solution - CL_{s+b} method



Frequentist solution - CL_{s+b} method

- Even though CL_{s+b} seems a valid method, it's has a problem

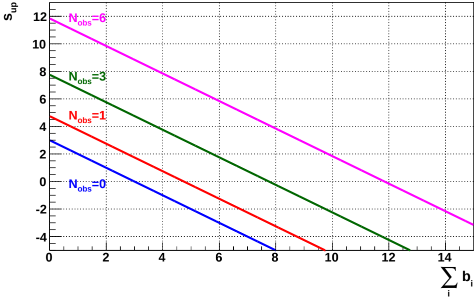


$\rightarrow$ downward fluctuation leads to $s_{\text{up}} < 0$!

- A fix could be to impose $s_{\text{up}} \geq 0$
  $\rightarrow$ Not satisfactory: all models predicting signal (even those predicting very small yield) are excluded with 5% probability

Frequentist solution - CL_{s+b} method

- Even though CL_{s+b} seems a valid method, it's has a problem



$\rightarrow$ downward fluctuation leads to $s_{\text{up}} < 0$!

- A fix could be to impose $s_{\text{up}} \geq 0$
  $\rightarrow$ Not satisfactory: all models predicting signal (even those predicting very small yield) are excluded with 5% probability

Frequentist solution - CL_s method

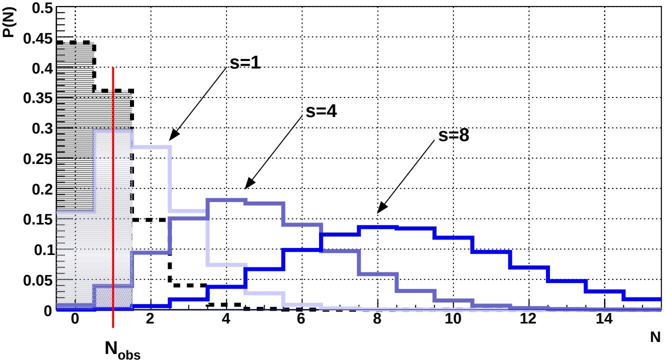
- Solution to previous issue: CL_s
- Instead of defining s_{up} by $CL_{s+b}(s_{\mathrm{up}}) = \alpha$, define it by

$$\boxed{CL_s(s_{\mathrm{up}}) = \alpha}$$

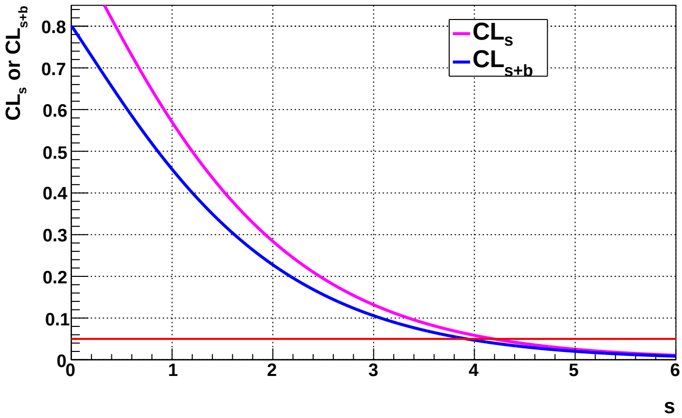
where $CL_s = \frac{CL_{s+b}}{CL_b}$, with

$$CL_b = \sum_{N=0}^{N_{\mathrm{obs}}} \frac{(\sum b_i)^N}{N!} e^{-\sum b_i}$$

Frequentist solution - CL_s method

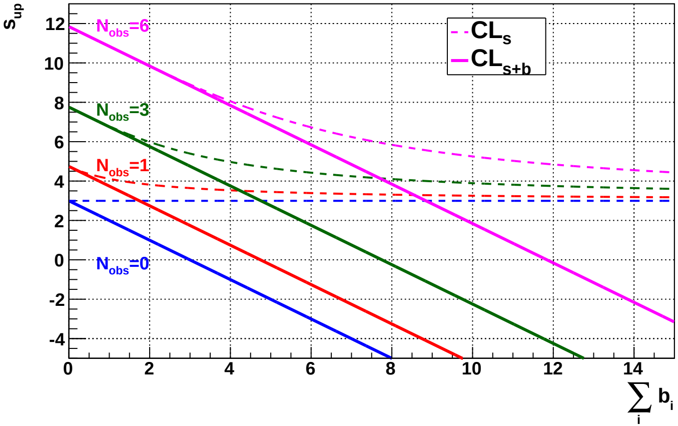


Frequentist solution - CL_s method



- Remark: upper limit always worse with CL_s (in this example: $s_{\text{up}} = 4.19$)

Frequentist solution - CL_s method



Frequentist solution

- Comments
  - Other solutions than CL_s exist to solve $s_{\text{up}} \leq 0$ problem (e.g. PCL)
  - CL_s is the current recommendation in ATLAS
  - Previous CL_s procedure is what CLsGenerator and McLimit do in single channel/no uncertainties case

Frequentist solution

- Comments
  - Other solutions than CL_s exist to solve $s_{\mathrm{up}} \leq 0$ problem (e.g. PCL)
  - CL_s is the current recommendation in ATLAS
  - Previous CL_s procedure is what CLsGenerator and McLimit do in single channel/no uncertainties case

Frequentist solution

- Comments
  - Other solutions than CL_s exist to solve $s_{\text{up}} \leq 0$ problem (e.g. PCL)
  - CL_s is the current recommendation in ATLAS
  - Previous CL_s procedure is what CLsGenerator and McLimit do in single channel/no uncertainties case

Bayesian solution

- Reminder: Bayes' theorem

$$P(A|B) = \frac{P(B|A)P(A)}{P(B)} \propto P(B|A)P(A)$$

$\rightarrow$ Bayesian methods use this theorem to make inference

- Bayes' theorem applied to previous example

$$\underbrace{f(s|N_{obs})}_{\text{posterior}} \propto \underbrace{P(N_{obs}|s)}_{\text{likelihood}} \underbrace{\pi(s)}_{\text{prior}} = \frac{\left(s+\sum_i b_i\right)^{N_{obs}}}{N_{obs}!} e^{-\left(s+\sum_i b_i\right)} \pi(s)$$

- Remark: philosophically very different from frequentist methods

$\rightarrow$ s **considered as a random variable**[2]

[2] hence the notation $P(N_{obs}|s) = P(N_{obs}; s)$

Bayesian solution

- Reminder: Bayes' theorem

$$P(A|B) = \frac{P(B|A)P(A)}{P(B)} \propto P(B|A)P(A)$$

$\rightarrow$ Bayesian methods use this theorem to make inference

- Bayes' theorem applied to previous example

$$\underbrace{f(s|N_{\text{obs}})}_{\text{posterior}} \propto \underbrace{P(N_{\text{obs}}|s)}_{\text{likelihood}} \underbrace{\pi(s)}_{\text{prior}} = \frac{\left(s+\sum_i b_i\right)^{N_{\text{obs}}}}{N_{\text{obs}}!} e^{-\left(s+\sum_i b_i\right)} \pi(s)$$

- Remark: philosophically very different from frequentist methods

$\rightarrow$ s **considered as a random variable**[2]

[2] hence the notation $P(N_{\text{obs}}|s) = P(N_{\text{obs}}; s)$

Bayesian solution

- Reminder: Bayes' theorem

$$P(A|B) = \frac{P(B|A)P(A)}{P(B)} \propto P(B|A)P(A)$$

$\rightarrow$ Bayesian methods use this theorem to make inference

- Bayes' theorem applied to previous example

$$\underbrace{f(s|N_{\text{obs}})}_{\text{posterior}} \propto \underbrace{P(N_{\text{obs}}|s)}_{\text{likelihood}} \underbrace{\pi(s)}_{\text{prior}} = \frac{\left(s+\sum_i b_i\right)^{N_{\text{obs}}}}{N_{\text{obs}}!} e^{-\left(s+\sum_i b_i\right)} \pi(s)$$

- Remark: philosophically very different from frequentist methods

$\rightarrow$ s **considered as a random variable**[2]

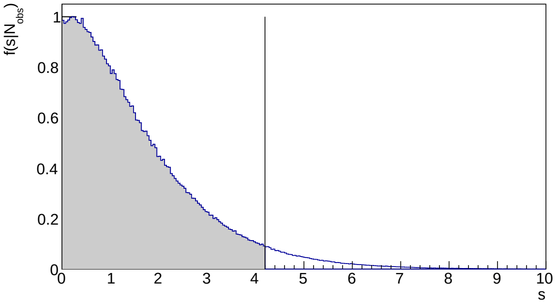
[2] hence the notation $P(N_{\text{obs}}|s) = P(N_{\text{obs}}; s)$

Bayesian solution

- Upper limit s_{up} defined by

$$\boxed{\int_0^{s_{\text{up}}} f(s|N_{\text{obs}})\text{d}s = 1-\alpha}$$

- For previous example ($N_{\text{obs}} = 1$, $\sum b_i = 0.82$) with uniform prior



$$\rightarrow s_{\text{up}} = 4.19$$

Frequentist Vs Bayesian

- CL_s and bayesian with uniform prior strictly equivalent in simple situation considered here

- Proof
  - $f(s|N_{\text{obs}}) = \frac{P(N_{\text{obs}}|s)}{\int_0^{\infty} P(N_{\text{obs}}|s)\text{ds}}$

    it can be shown that $\int_0^{\infty} P(N_{\text{obs}}|s)ds = CL_b$
  - $1-\alpha = \int_0^{s_{\text{up}}} f(s|N_{\text{obs}})\text{ds} = \frac{1}{CL_b}\int_0^{s_{\text{up}}} P(N_{\text{obs}}|s)\text{ds}$

    it can be shown that $\int_0^{s_{\text{up}}} P(N_{\text{obs}}|s)\text{ds} = CL_b - CL_{s+b}(s_{\text{up}})$

    $$\Rightarrow \alpha = \frac{CL_{s+b}(s_{\text{up}})}{CL_b} = CL_s(s_{\text{up}})$$

Frequentist Vs Bayesian

- CL_s and bayesian with uniform prior strictly equivalent in simple situation considered here

- Proof
  - $f(s|N_{\text{obs}}) = \frac{P(N_{\text{obs}}|s)}{\int_0^\infty P(N_{\text{obs}}|s)\text{ds}}$

    it can be shown that $\int_0^\infty P(N_{\text{obs}}|s)ds = CL_b$
  - $1-\alpha = \int_0^{s_{\text{up}}} f(s|N_{\text{obs}})\text{ds} = \frac{1}{CL_b}\int_0^{s_{\text{up}}} P(N_{\text{obs}}|s)\text{ds}$

    it can be shown that $\int_0^{s_{\text{up}}} P(N_{\text{obs}}|s)\text{ds} = CL_b - CL_{s+b}(s_{\text{up}})$

$$\Rightarrow \alpha = \frac{CL_{s+b}(s_{\text{up}})}{CL_b} = CL_s(s_{\text{up}})$$

Analytical solution

- No need to run sophisticated tools to compute previous limits

$\rightarrow$ Analytical solutions exist

- Analytical solutions (proofs on next slide)
  - CL_{s+b}:

$$\boxed{s_{\mathbf{up}} = 0.5 \times F^{-1}_{\chi^2}(1-\alpha; 2(N_{\mathbf{obs}}+1)) - \sum b_i}$$

`0.5*ROOT::Math::chisquared_quantile(1-`α`,2*(`N_{obs}`+1))-`b

  - CL_s and bayesian with uniform prior :

$$\boxed{s_{\mathbf{up}} = 0.5 \times F^{-1}_{\chi^2}\left(1-\alpha\left[1-F_{\chi^2}\left(2\sum b_i; 2(N_{\mathbf{obs}}+1)\right)\right]; 2(N_{\mathbf{obs}}+1)\right) - \sum b_i}$$

`0.5*ROOT::Math::chisquared_quantile(1-`α`*(1-`

`ROOT::Math::chisquared_cdf(2*`b`,2*(`N_{obs}`+1))),2*(`N_{obs}`+1))-`b

Analytical solution

- No need to run sophisticated tools to compute previous limits

  $\rightarrow$ Analytical solutions exist

- Analytical solutions (proofs on next slide)
  - CL_{s+b}:

$$\boxed{s_{\mathbf{up}} = 0.5 \times F^{-1}_{\chi^2}\left(1-\alpha; 2\left(N_{\mathbf{obs}}+1\right)\right) - \sum b_i}$$

`0.5*ROOT::Math::chisquared_quantile(1-`α`,2*(`N_{obs}`+1))-`b

  - CL_s and bayesian with uniform prior :

$$\boxed{s_{\mathbf{up}} = 0.5 \times F^{-1}_{\chi^2}\left(1-\alpha\left[1-F_{\chi^2}\left(2\sum b_i; 2\left(N_{\mathbf{obs}}+1\right)\right)\right]; 2\left(N_{\mathbf{obs}}+1\right)\right) - \sum b_i}$$

`0.5*ROOT::Math::chisquared_quantile(1-`α`*(1-`
`ROOT::Math::chisquared_cdf(2*`b`,2*(`N_{obs}`+1))),2*(`N_{obs}`+1))-`b

Analytical solution

- Demonstration of previous relations
  - Consider first a $CL_{\mathrm{s+b}}$ limit. Upper limit μ_{up} at $1-\alpha$ confidence level given by

$$CL_{\mathrm{s+b}} = \sum_{n=0}^{N\mathrm{obs}} \mathscr{P}(N; s + \sum b_i) = \alpha$$

    Poisson c. d. f. given by

$$\sum_{n=0}^{N_{\mathrm{obs}}} \mathscr{P}(N; s + \sum b_i) = 1 - F_{\chi^2}(2(s + \sum b_i); 2(N_{\mathrm{obs}} + 1))$$

    Thus

$$s_{\mathrm{up}} = 0.5 \times F^{-1}_{\chi^2}(1-\alpha; 2(N_{\mathrm{obs}}+1)) - \sum b_i$$

  - For CL_s, just need to replace α by $\alpha \times CL_{\mathrm{b}}$ (since CL_b is independant of signal)

Remark

- Note that previous CL_{s+b} formula is the one used in SS analysis to set upper limit for samples with yield=0

$$\text{Yield}_{\text{up}} = 0.5 \times F_{\chi^2}^{-1}\left(0.68; 2\left(0+1\right)\right) \simeq 1.14$$

Summary of methods discussed so far

	Method	Comment
Frequentist	CL_{s+b}	unphysical when downward fluctuation
	CL_s	equivalent to bayesian w/ uniform prior
	PCL	maybe solution of the future
Bayesian	Bayesian (uniform prior)	equivalent to CL_s
	Bayesian (other priors)	

$\rightarrow$ Quite a lot of solutions for such a simple problem

It's just the beginning !

Summary of methods discussed so far

	Method	Comment
Frequentist	CL_{s+b}	unphysical when downward fluctuation
	CL_s	equivalent to bayesian w/ uniform prior
	PCL	maybe solution of the future
Bayesian	Bayesian (uniform prior)	equivalent to CL_s
	Bayesian (other priors)	

$\rightarrow$ Quite a lot of solutions for such a simple problem

It's just the beginning !

Sensitivity of an analysis

- Sensitivity of an analysis not given by s_{up} but by some quantity measuring by how much s_{up} differs from some nominal signal expectation s_{nom}
  - Example: SUSY predicts s_{nom} sgluon events after selection
    - if $s_{\text{up}} < s_{\text{nom}}$: sgluon excluded
    - if $s_{\text{up}} \geq s_{\text{nom}}$: sgluon not excluded
- Sensitivity usually given by

$$\mu_{\text{up}} = \frac{s_{\text{up}}}{s_{\text{nom}}}$$

If $\mu_{\text{up}} \gg 1$, analysis not sensitive

Sensitivity of an analysis

- Sensitivity of an analysis not given by s_{up} but by some quantity measuring by how much s_{up} differs from some nominal signal expectation s_{nom}
  - Example: SUSY predicts s_{nom} sgluon events after selection
    - if $s_{\text{up}} < s_{\text{nom}}$: sgluon excluded
    - if $s_{\text{up}} \geq s_{\text{nom}}$: sgluon not excluded
- Sensitivity usually given by

$$\mu_{\text{up}} = \frac{s_{\text{up}}}{s_{\text{nom}}}$$

If $\mu_{\text{up}} \gg 1$, analysis not sensitive

Comments on historical $\mu_{\text{up}} = 1.64\sqrt{b}/s_{\text{nom}}$

- In order to compute sensitivities fast, people sometimes use

$$\mu_{\text{up}} = 1.64 \frac{\sqrt{b}}{s_{\text{nom}}} \qquad (b : \text{total background})$$

- From this formula one often predicts that sensitivity should scale as $1/\sqrt{\text{Lumi}}$

→ Where does this come from ?
→ What approximations are behind ?
→ Can we do better ?
→ Is the $1/\sqrt{\text{Lumi}}$ scaling correct ?

Comments on historical $\mu_{\text{up}} = 1.64\sqrt{b}/s_{\text{nom}}$

- In order to compute sensitivities fast, people sometimes use

$$\mu_{\text{up}} = 1.64\frac{\sqrt{b}}{s_{\text{nom}}} \qquad (b : \text{total background})$$

- From this formula one often predicts that sensitivity should scale as $1/\sqrt{\text{Lumi}}$

→ Where does this come from ?
→ What approximations are behind ?
→ Can we do better ?
→ Is the $1/\sqrt{\text{Lumi}}$ scaling correct ?

Gaussian approximation

- Analytical solution in gaussian approximation
  - When $s + \sum b_i$ is large, poisson $\simeq$ gaussian with $\sigma = \sqrt{s + \sum b_i}$
  - We can show that, for CL_{s+b} (see next slide)

$$\boxed{s_{\mathrm{up}} = N_{\mathrm{obs}} - \sum b_i + \frac{\left[\Phi^{-1}(1-\alpha)\right]^2}{2}\left[1 + \sqrt{1 + 4\frac{N_{\mathrm{obs}}}{\left[\Phi^{-1}(1-\alpha)\right]^2}}\right]}$$

where Φ is the c. d. f. of the standard normal distribution

Gaussian approximation

- Analytical solution in gaussian approximation
  - When $s+\sum b_i$ is large, poisson $\simeq$ gaussian with $\sigma = \sqrt{s+\sum b_i}$
  - We can show that, for CL_{s+b} (see next slide)

$$\boxed{s_{\mathrm{up}} = N_{\mathrm{obs}} - \sum b_i + \frac{\left[\Phi^{-1}(1-\alpha)\right]^2}{2}\left[1+\sqrt{1+4\frac{N_{\mathrm{obs}}}{\left[\Phi^{-1}(1-\alpha)\right]^2}}\right]}$$

where Φ is the c. d. f. of the standard normal distribution

Gaussian approximation - Demonstration

- For CL_{s+b}, we have by definition

$$CL_{\mathrm{s+b}} = \int_0^{N_{\mathrm{obs}}} G\left(N; s_{\mathrm{up}} + \sum b_i, \sqrt{s_{\mathrm{up}} + \sum b_i}\right) \mathrm{d}N = \alpha$$

$$\Rightarrow \int_0^{\frac{N_{\mathrm{obs}} - (s_{\mathrm{up}} + \sum b_i)}{\sqrt{s_{\mathrm{up}} + \sum b_i}}} G(N; 0, 1) \mathrm{d}N = \alpha$$

$$\Rightarrow \Phi\left(\frac{N_{\mathrm{obs}} - (s_{\mathrm{up}} + \sum b_i)}{\sqrt{s_{\mathrm{up}} + \sum b_i}}\right) = \alpha$$

$$\Rightarrow N_{\mathrm{obs}} - (s_{\mathrm{up}} + \sum b_i) = -\sqrt{s_{\mathrm{up}} + \sum b_i} \times \Phi^{-1}(1 - \alpha)$$

  Solving for s_{up} gives expression on previous slide

- For CL_s, just need to replace α by $\alpha \times \underbrace{\Phi\left(\frac{N_{\mathrm{obs}-\sum b_i}}{\sqrt{\sum b_i}}\right)}_{CL_b}$ (as for Poisson)

Comments on historical $\mu_{up} = 1.64\sqrt{b}/s_{nom}$

$$\boxed{\mu_{up} = \frac{N_{obs} - \sum b_i}{s_{nom}} + \frac{\left[\Phi^{-1}(1-\alpha)\right]^2}{2s_{nom}}\left[1 + \sqrt{1 + 4\frac{N_{obs}}{[\Phi^{-1}(1-\alpha)]^2}}\right]} \quad (\star)$$

- Let's assume that
  1. $N_{obs} = \sum b_i$ (i.e. expected limit)
  2. $\sum b_i \gg \left[\Phi^{-1}(1-\alpha)\right]^2/4$
     - note that $\Phi^{-1}(1-\alpha) = 1.64$ for $\alpha = 0.05$, corresponding to $\sum b_i \gg 0.7$
     - this assumption corresponds more or less to the gaussian approx.

  Then, one sees directly from $(\star)$ that

$$\boxed{\mu_{up} \simeq \Phi^{-1}(1-\alpha)\frac{\sqrt{\sum b_i}}{s_{nom}} = 1.64\frac{\sqrt{\sum b_i}}{s_{nom}}}$$

Comments on historical $\mu_{\text{up}} = 1.64\sqrt{b}/s_{\text{nom}}$

$$\boxed{\mu_{\text{up}} = \frac{N_{\text{obs}} - \sum b_i}{s_{\text{nom}}} + \frac{\left[\Phi^{-1}(1-\alpha)\right]^2}{2s_{\text{nom}}}\left[1 + \sqrt{1 + 4\frac{N_{\text{obs}}}{\left[\Phi^{-1}(1-\alpha)\right]^2}}\right]} \qquad (\star)$$

- Let's assume that
  1. $N_{\text{obs}} = \sum b_i$ (i.e. expected limit)
  2. $\sum b_i \gg \left[\Phi^{-1}(1-\alpha)\right]^2/4$
     - note that $\Phi^{-1}(1-\alpha) = 1.64$ for $\alpha = 0.05$, corresponding to $\sum b_i \gg 0.7$
     - this assumption corresponds more or less to the gaussian approx.

Then, one sees directly from $(\star)$ that

$$\boxed{\mu_{\text{up}} \simeq \Phi^{-1}(1-\alpha)\frac{\sqrt{\sum b_i}}{s_{\text{nom}}} = 1.64\frac{\sqrt{\sum b_i}}{s_{\text{nom}}}}$$

Comments on historical $\mu_{\text{up}} = 1.64\sqrt{b}/s_{\text{nom}}$

- Where does this come from ? What approximations are behind ?
  $\rightarrow$ CL_{s+b} method under gaussian approximation

- Can we do better ?
  $\rightarrow$ Yes: use analytical CL_s poisson solution (no more complicated to use and more correct)

$$\boxed{s_{\text{up}} = 0.5 \times F^{-1}_{\chi^2}\left(1 - \alpha\left[1 - F_{\chi^2}\left(2\sum b_i; 2\left(N_{\text{obs}}+1\right)\right)\right]; 2\left(N_{\text{obs}}+1\right)\right) - \sum b_i}$$

0.5*ROOT::Math::chisquared_quantile(1-α*(1-ROOT::Math::chisquared_cdf(2*b,2*(N_{obs}+1))),2*(N_{obs}+1))-b

Comments on historical $\mu_{\text{up}} = 1.64\sqrt{b}/s_{\text{nom}}$

- Where does this come from ? What approximations are behind ?
  $\rightarrow$ CL_{s+b} method under gaussian approximation

- Can we do better ?
  $\rightarrow$ Yes: use analytical CL_s poisson solution (no more complicated to use and more correct)

$$\boxed{s_{\text{up}} = 0.5 \times F^{-1}_{\chi^2}\left(1-\alpha\left[1-F_{\chi^2}\left(2\sum b_i; 2\left(N_{\text{obs}}+1\right)\right)\right]; 2\left(N_{\text{obs}}+1\right)\right) - \sum b_i}$$

```
0.5*ROOT::Math::chisquared_quantile(1-α*(1-
        ROOT::Math::chisquared_cdf(2*b,2*(N_obs+1))),2*(N_obs+1))-b
```

Comments on historical $\mu_{\text{up}} = 1.64\sqrt{b}/s_{\text{nom}}$

- Where does this come from ? What approximations are behind ?
  $\rightarrow$ CL_{s+b} method under gaussian approximation

- Can we do better ?
  $\rightarrow$ Yes: use analytical CL_s poisson solution (no more complicated to use and more correct)

$$\boxed{s_{\text{up}} = 0.5 \times F^{-1}_{\chi^2}\left(1 - \alpha\left[1 - F_{\chi^2}\left(2\sum b_i; 2\left(N_{\textbf{obs}}+1\right)\right)\right]; 2\left(N_{\textbf{obs}}+1\right)\right) - \sum b_i}$$

```
0.5*ROOT::Math::chisquared_quantile(1-α*(1-
        ROOT::Math::chisquared_cdf(2*b,2*(N_obs+1))),2*(N_obs+1))-b
```

Comments on historical $\mu_{\text{up}} = 1.64\sqrt{b}/s_{\text{nom}}$

- Where does this come from ? What approximations are behind ?
  $\rightarrow$ CL_{s+b} method under gaussian approximation

- Can we do better ?
  $\rightarrow$ Yes: use analytical CL_s poisson solution (no more complicated to use and more correct)

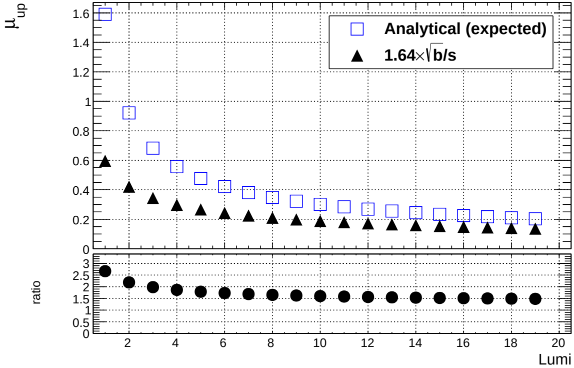
$$\boxed{s_{\mathbf{up}} = 0.5 \times F^{-1}_{\chi^2}\left(1 - \alpha\left[1 - F_{\chi^2}\left(2\sum b_i; 2\left(N_{\mathbf{obs}} + 1\right)\right)\right]; 2\left(N_{\mathbf{obs}} + 1\right)\right) - \sum b_i}$$

0.5*ROOT::Math::chisquared_quantile(1-α*(1-

ROOT::Math::chisquared_cdf(2*b,2*(N_{obs}+1))),2*(N_{obs}+1))-b

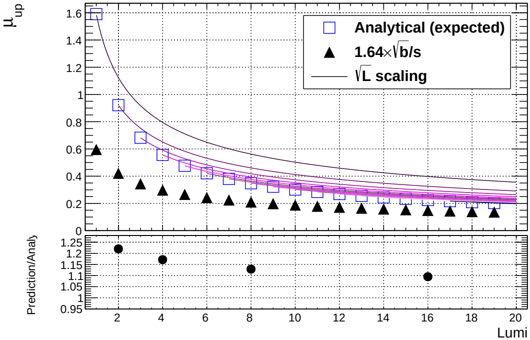
Comments on historical $\mu_{\rm up} = 1.64\sqrt{b}/s_{\rm nom}$

- Consider example with $b = 0.82$, $s_{\rm nom} = 2.49$ for Lumi$= 1$



Comments on historical $\mu_{up} = 1.64\sqrt{b}/s_{nom}$

- Is the $1/\sqrt{\text{Lumi}}$ scaling correct ?
  $\rightarrow$ As before: $b = 0.82$, $s_{nom} = 2.49$ for Lumi$= 1$



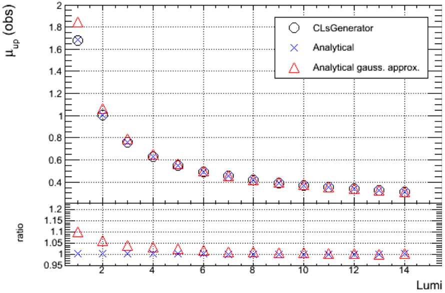
$\Rightarrow$ $1/\sqrt{\text{Lumi}}$ scaling rather good approximation when number of events not too small

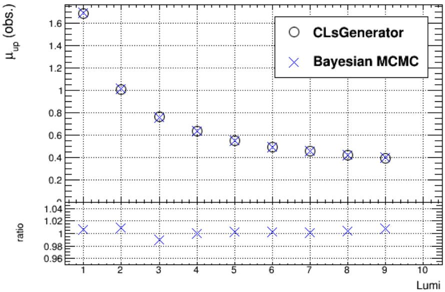
Softwares

- No softwares needed, use analytical (unless you want bayesian w/ non-uniform prior or profile likelihood)

Validation of CLsGenerator and BayesianMCMC

- Use analytical solution to validate CLsGenerator and BayesianMCMC





→ CLsGenerator and BayesianMCMC validated in single channel - no uncertainties case

Multiple channels - No uncertainties

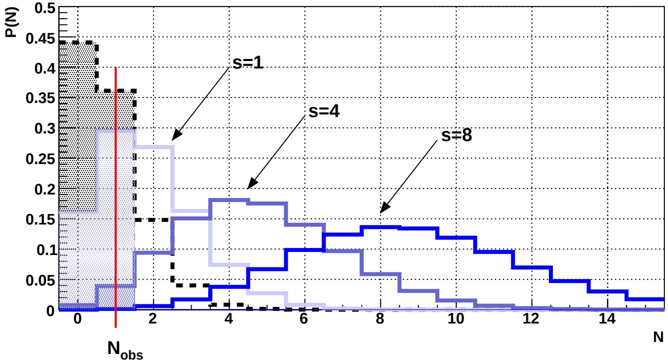
Generalizing frequentist methods

- Methods presented in simple "Single channel - No uncertainties" case can be used as prototypes for more general cases
- What we did previously was to
  1. Pick variable with discrimating power between background and signal+background hypothesis
  2. Determine distribution of this variable
  3. Compute $p-values$ as a function of signal yield s
  4. Determine upper limit s_{up}
- All frequentist methods based on this general scheme

Generalizing frequentist methods

- Methods presented in simple "Single channel - No uncertainties" case can be used as prototypes for more general cases
- What we did previously was to
  1. Pick variable with discrimating power between background and signal+background hypothesis
  2. Determine distribution of this variable
  3. Compute $p - values$ as a function of signal yield s
  4. Determine upper limit s_{up}
- All frequentist methods based on this general scheme

Generalizing frequentist methods



Generalizing frequentist methods - Test statistic

- How to choose variable ?
  - Need single discriminating variable summarizing information from all channels

    → Can't use observed yield anymore

    → Use likelihood based variables

$$\boxed{\mathscr{L}_{\text{joint}} = \prod_{\text{c:channels}} \mathscr{L}_{\text{c}}}$$

  - Discriminating power as good as possible

    → Likelihood ratios are a good choice (Neyman-Pearson Lemma)

- Terminology: this variable is called test statistic or simply test

Generalizing frequentist methods - Test statistic

- How to choose variable ?
  - Need single discriminating variable summarizing information from all channels

    → Can't use observed yield anymore

    → Use likelihood based variables

$$\boxed{\mathscr{L}_{\text{joint}} = \prod_{\text{c:channels}} \mathscr{L}_{\text{c}}}$$

  - Discriminating power as good as possible

    → Likelihood ratios are a good choice (Neyman-Pearson Lemma)

- Terminology: this variable is called test statistic or simply test

Generalizing frequentist methods - Test statistic

- How to choose variable ?
  - Need single discriminating variable summarizing information from all channels

    $\rightarrow$ Can't use observed yield anymore

    $\rightarrow$ Use likelihood based variables

$$\boxed{\mathscr{L}_{\text{joint}} = \prod_{\text{c:channels}} \mathscr{L}_{\text{c}}}$$

  - Discriminating power as good as possible

    $\rightarrow$ Likelihood ratios are a good choice (Neyman-Pearson Lemma)

- Terminology: this variable is called test statistic or simply test

Generalizing frequentist methods - Test statistic

- How to choose variable ?
  - Need single discriminating variable summarizing information from all channels

    $\rightarrow$ Can't use observed yield anymore

    $\rightarrow$ Use likelihood based variables

$$\boxed{\mathscr{L}_{\text{joint}} = \prod_{c:\text{channels}} \mathscr{L}_{\text{c}}}$$

  - Discriminating power as good as possible

    $\rightarrow$ Likelihood ratios are a good choice (Neyman-Pearson Lemma)

- Terminology: this variable is called test statistic or simply test

Generalizing frequentist methods - Test statistic

- How to choose variable ?
  - Need single discriminating variable summarizing information from all channels

    → Can't use observed yield anymore

    → Use likelihood based variables

$$\boxed{\mathscr{L}_{\text{joint}} = \prod_{c\text{:channels}} \mathscr{L}_{\text{c}}}$$

  - Discriminating power as good as possible

    → Likelihood ratios are a good choice (Neyman-Pearson Lemma)

- Terminology: this variable is called test statistic or simply test

Generalizing frequentist methods - Test statistic

- How to choose variable ?
  - Need single discriminating variable summarizing information from all channels

    $\rightarrow$ Can't use observed yield anymore

    $\rightarrow$ Use likelihood based variables

$$\boxed{\mathscr{L}_{\text{joint}} = \prod_{c:\text{channels}} \mathscr{L}_{\text{c}}}$$

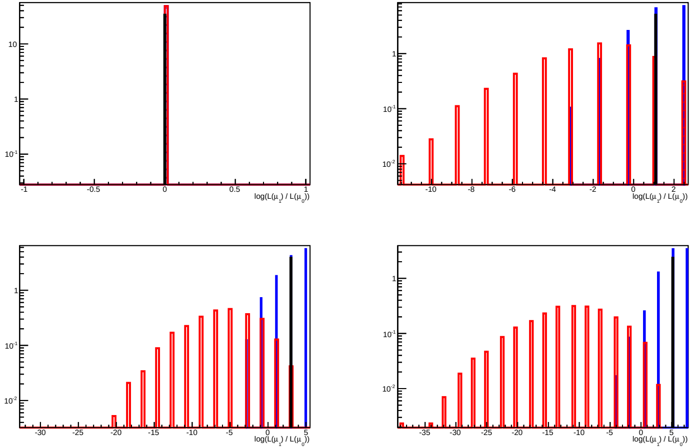
  - Discriminating power as good as possible

    $\rightarrow$ Likelihood ratios are a good choice (Neyman-Pearson Lemma)

- Terminology: this variable is called test statistic or simply test

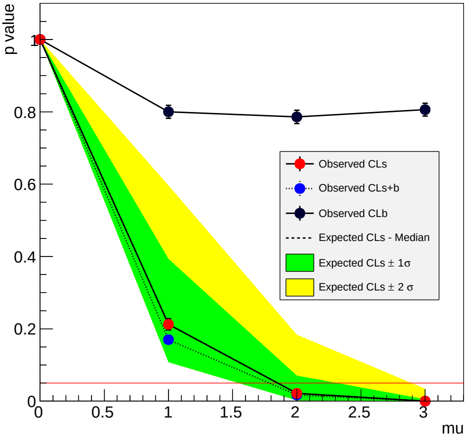
Generalizing frequentist methods - Test statistic

- Once a test is chosen, apply the rest of the procedure as before



Generalizing frequentist methods - Test statistic

- Once a test is chosen, apply the rest of the procedure as before



Common test statistics

Test statistic	Comment	Software
$q_\mu = -2\ln\frac{\mathscr{L}(\mu)}{\mathscr{L}(\mu=0)}$ (simple likelihood ratio)	Used at LEP, Tevatron, LHC	McLimit, CLsGenerator, RooStats
$q'_\mu = -2\ln\frac{\mathscr{L}(\mu)}{\mathscr{L}(\hat{\mu})}$ (profile likelihood ratio)	Used at LHC	RooStats

- Difference w. r. t. single channel case: test depends on μ
  - Distribution under background hypothesis depends on μ
  - q_μ^{obs} depends on μ

CLsGenerator/McLimit test statistic

- Number of channels: n_c
- Test

$$q_\mu = -2\ln \frac{\mathscr{L}_1(\mu)\times\mathscr{L}_2(\mu)\times\ldots\ \times\mathscr{L}_{n_c}(\mu)}{\mathscr{L}_1(\mu{=}0)\times\mathscr{L}_2(\mu{=}0)\times\ldots\ \times\mathscr{L}_{n_c}(\mu{=}0)}$$

$$\Rightarrow \boxed{q_\mu = \sum_{c=1}^{n_c} q_\mu^c}$$

- For each channel

$$q_\mu^c = -2\ln\left[\frac{(\mu s_{\text{nom}}+\sum b_i)^{N_{\text{obs}}}\, e^{-(\mu s_{\text{nom}}+\sum b_i)}}{(\sum b_i)^{N_{\text{obs}}}\, e^{-\sum b_i}}\right]$$

$$\boxed{q_\mu^c = 2\left(\mu s_{\text{nom}} - N_{\text{obs}}\ln\left(\frac{\mu s_{\text{nom}}+\sum b_i}{\sum b_i}\right)\right)}$$

CLsGenerator/McLimit test statistic

- In single channel case, using q_μ or observed yield as test is equivalent
  - Everything we said for single channel case remains valid

Generalizing bayesian method

- Reminder

$$\text{posterior} \propto \text{likelihood} \times \text{prior}$$

$$\int_0^{s_{\text{up}}} \text{posterior} = 1 - \alpha$$

- Multiple channels: use joint likelihood

$$\boxed{\mathscr{L}_{\text{joint}} = \prod_{c:\text{channels}} \mathscr{L}_{\text{c}}}$$

Softwares

- CLsGenerator and BayesianMCMC can be used for multiple channels

Validation of CLsGenerator and BayesianMCMC

- We expect that increasing luminosity by factor N is equivalent to increasing number of channels (having the same sensitivity) by the same factor (see next two slides)
- Here we check that CLsGenerator and BayesianMCMC are in agreement with this expectation
  - Validates in both cases implementation of channel combination

Demonstration equivalence lumi/#channels (CLsGenerator)

- Single channel: $q_\mu = 2\left[\mu s - N_{\text{obs}} \ln \frac{\mu s + b}{b}\right]$
- N channels: $q_{\mu,1} = 2\left[\mu \sum_{c:channels} s_c - \sum_{c:channels} N^c_{\text{obs}} \ln \frac{\mu s_c + b_c}{b_c}\right]$
  - Assume yields are the same in all channels ($s_c = s$, $b_c = b$, $N^c_{\text{obs}} = N_{\text{obs}}$):
    $q_{\mu,1} = 2\left[\mu N s - N_{\text{obs}} \ln\left(\prod_{c:channels} \frac{\mu s_c + b_c}{b_c}\right)\right] = N \times 2\left[\mu s - N_{\text{obs}} \ln \frac{\mu s + b}{b}\right]$
    $\Rightarrow q_{\mu,1} = N q_\mu$
- Multiplying lumi. by N: $q_{\mu,2} = 2\left[\mu s' - N'_{\text{obs}} \ln \frac{\mu s' + b'}{b'}\right]$
  - $s' = Ns$, $b' = Nb$ and $N'_{\text{obs}} = N N_{\text{obs}}$

$\Rightarrow q_{\mu,2} = N \times 2\left[\mu s - N_{\text{obs}} \ln \frac{\mu s + b}{b}\right] = N q_\mu$

- Conclusion: $q_{\mu,1} = q_{\mu,2}$

Demonstration equivalence lumi/#channels (BayesianMCMC)

- Single channel: $\mathcal{L}(\mu) = \frac{(\mu s+b)^{N_{\text{obs}}}}{N_{\text{obs}}\,!} e^{-(\mu s+b)}$
- N channels:
  $\mathcal{L}_1(\mu) = \prod_{c:channels} \mathcal{L}_c(\mu) = \prod_{c:channels} \frac{(\mu s_c+b_c)^{N^c_{\text{obs}}}}{N^c_{\text{obs}}\,!} e^{-(\mu s_c+b_c)}$
  - Assume yields are the same in all channels ($s_c = s$, $b_c = b$, $N^c_{\text{obs}} = N_{\text{obs}}$):
    $\Rightarrow \mathcal{L}_1(\mu) = \left[\frac{(\mu s+b)^{N_{\text{obs}}}}{N_{\text{obs}}\,!} e^{-(\mu s+b)}\right]^N = \frac{1}{(N_{\text{obs}}!)^N} (\mu s + b)^{NN_{\text{obs}}} e^{-N(\mu s+b)}$
- Multiplying lumi. by N: $\mathcal{L}_2(\mu) = \frac{(\mu s'+b')^{N'_{\text{obs}}}}{N'_{\text{obs}}\,!} e^{-(\mu s'+b')}$
  - $s' = Ns$, $b' = Nb$ and $N'_{\text{obs}} = NN_{\text{obs}}$

$$\Rightarrow \mathcal{L}_2(\mu) = \frac{(\mu Ns+Nb)^{NN_{\text{obs}}}}{(NN_{\text{obs}})!} e^{-(\mu Ns+Nb)} = \frac{N^{NN_{\text{obs}}}}{(NN_{\text{obs}})!} (\mu s + b)^{NN_{\text{obs}}} e^{-N(\mu s+b)}$$

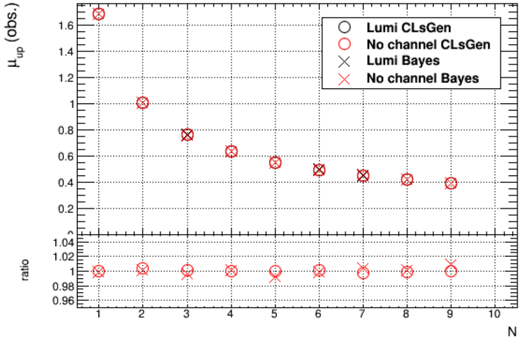
- Conclusion: $\mathcal{L}_1(\mu) \propto \mathcal{L}_2(\mu)$ so inference on μ is the same

Validation of CLsGenerator and BayesianMCMC

- As always, consider
  - $N_{obs} = 1$
  - $b = 0.82$
  - $s_{\text{nom}} = 2.49$
- Compare two limits

first one computed from one channel with – $N_{obs} = 1 \times N$ – $b = 0.82 \times N$ – $s_{\text{nom}} = 2.49 \times N$	second one computed from N identical channels with – $N_{obs} = 1$ – $b = 0.82$ – $s_{\text{nom}} = 2.49$

Validation of CLsGenerator and BayesianMCMC



→ Expectation verified in both CLsGenerator and BayesianMCMC

Uncertainties

Problematic

- Nominal signal and background yields (s_{nom} and b_i) not known with infinite precision
- Two types of uncertainties
  - Statistical (finite sample size)
  - Systematic (e.g. JES, JVF, etc.)

$\rightarrow$ How to account for uncertainties in limit setting ?

Statistical uncertainties

- Consider treatment of stat. uncert. as done in McLimit, CLsGenerator and BayesianMCMC (HistFactory makes things differently)

- Signal and background are considered as random variables with
  - average = nominal yields ($\sum w_i$)
  - standard deviation $= \begin{cases} \text{stat. uncert. } (\sqrt{\sum w_i^2}) \text{ when nominal yield} \neq 0 \\ \text{upper limit (1.14 lumi rescaled) when yield} = 0 \end{cases}$

Single channel

$$\mathscr{L}(\mu, s', \{b'_i\}) = \frac{(\mu s' + \sum_i b'_i)^{N_{\text{obs}}}}{N_{\text{obs}}!} e^{-(\mu s' + \sum_i b'_i)} \underbrace{f(s'; s_{\text{nom}}, \sigma_s)}_{\text{constraint signal}} \prod_i \underbrace{f(b'_i; b_i, \sigma_i)}_{\text{constraint background i}}$$

- Terminology: s' and $\{b'_i\}$ are called nuisance parameters

Statistical uncertainties

- Consider treatment of stat. uncert. as done in McLimit, CLsGenerator and BayesianMCMC (HistFactory makes things differently)

- Signal and background are considered as random variables with
  - average = nominal yields ($\sum w_i$)
  - standard deviation $= \begin{cases} \text{stat. uncert. } (\sqrt{\sum w_i^2}) \text{ when nominal yield} \neq 0 \\ \text{upper limit (1.14 lumi rescaled) when yield} = 0 \end{cases}$

Single channel

$$\mathscr{L}(\mu, s', \{b_i'\}) = \frac{(\mu s' + \sum_i b_i')^{N_{\text{obs}}}}{N_{\text{obs}}!} e^{-(\mu s' + \sum_i b_i')} \underbrace{f(s'; s_{\text{nom}}, \sigma_s)}_{\text{constraint signal}} \prod_i \underbrace{f(b_i'; b_i, \sigma_i)}_{\text{constraint background i}}$$

- Terminology: s' and $\{b_i'\}$ are called nuisance parameters

Statistical uncertainties

- Consider treatment of stat. uncert. as done in McLimit, CLsGenerator and BayesianMCMC (HistFactory makes things differently)

- Signal and background are considered as random variables with
  - average = nominal yields ($\sum w_i$)
  - standard deviation $= \begin{cases} \text{stat. uncert. } (\sqrt{\sum w_i^2}) \text{ when nominal yield} \neq 0 \\ \text{upper limit (1.14 lumi rescaled) when yield} = 0 \end{cases}$

Single channel

$$\boxed{\mathscr{L}(\mu, s', \{b'_i\}) = \frac{(\mu s' + \sum_i b'_i)^{N_{\text{obs}}}}{N_{\text{obs}}!} e^{-(\mu s' + \sum_i b'_i)} \underbrace{f(s'; s_{\text{nom}}, \sigma_s)}_{\text{constraint signal}} \prod_i \underbrace{f(b'_i; b_i, \sigma_i)}_{\text{constraint background i}}}$$

- Terminology: s' and $\{b'_i\}$ are called nuisance parameters

What McLimit/CLsGenerator/BayesianMCMC do ?

Single channel

$$\mathscr{L}(\mu, s', \{b'_i\}) = \frac{(\mu s' + \sum_i b'_i)^{N_{\text{obs}}}}{N_{\text{obs}}!} e^{-(\mu s' + \sum_i b'_i)} \underbrace{f(s'; s_{\text{nom}}, \sigma_s)}_{\text{constraint signal}} \prod_i \underbrace{f(b'_i; b_i, \sigma_i)}_{\text{constraint background i}}$$

- McLimit/CLsGenerator toss pseudo-experiments in which
  - N_{obs} is sampled from marginal likelihood (hence the hybrid nature of the method)

$$\mathscr{L}(\mu) = \int \frac{(\mu s' + \sum_i b'_i)^{N_{\text{obs}}}}{N_{\text{obs}}!} e^{-(\mu s' + \sum_i b'_i)} f(s'; s_{\text{nom}}, \sigma_s) \prod_i f(b'_i; b_i, \sigma_i) \mathrm{d}s' \prod_i \mathrm{d}b'_i \quad (\square)$$

  - Test is computed using nominal values of nuisance parameters

$$\mathscr{L}(\mu, s' = s_{nom}, \{b'_i\} = \{b_i\}) = \frac{(\mu s_{\text{nom}} + \sum_i b_i)^{N_{\text{obs}}}}{N_{\text{obs}}!} e^{-(\mu s_{\text{nom}} + \sum_i b_i)}$$

- BayesianMCMC uses marginal likelihood ($\square$) as posterior (integration done by Markov Chain Monte Carlo)

What McLimit/CLsGenerator/BayesianMCMC do ?

Single channel

$$\mathscr{L}(\mu, s', \{b'_i\}) = \frac{(\mu s' + \sum_i b'_i)^{N_{\text{obs}}}}{N_{\text{obs}}!} e^{-(\mu s' + \sum_i b'_i)} \underbrace{f(s'; s_{\text{nom}}, \sigma_s)}_{\text{constraint signal}} \prod_i \underbrace{f(b'_i; b_i, \sigma_i)}_{\text{constraint background i}}$$

- McLimit/CLsGenerator toss pseudo-experiments in which
  - N_{obs} is sampled from marginal likelihood (hence the hybrid nature of the method)

$$\mathscr{L}(\mu) = \int \frac{(\mu s' + \sum_i b'_i)^{N_{\text{obs}}}}{N_{\text{obs}}!} e^{-(\mu s' + \sum_i b'_i)} f(s'; s_{\text{nom}}, \sigma_s) \prod_i f(b'_i; b_i, \sigma_i) \mathrm{d}s' \prod_i \mathrm{d}b'_i \quad (\square)$$

  - Test is computed using nominal values of nuisance parameters

$$\mathscr{L}(\mu, s' = s_{nom}, \{b'_i\} = \{b_i\}) = \frac{(\mu s_{\text{nom}} + \sum_i b_i)^{N_{\text{obs}}}}{N_{\text{obs}}!} e^{-(\mu s_{\text{nom}} + \sum_i b_i)}$$

- BayesianMCMC uses marginal likelihood ($\square$) as posterior (integration done by Markov Chain Monte Carlo)

What McLimit/CLsGenerator/BayesianMCMC do ?

Single channel

$$\mathscr{L}(\mu, s', \{b'_i\}) = \frac{(\mu s' + \sum_i b'_i)^{N_{\text{obs}}}}{N_{\text{obs}}!} e^{-(\mu s' + \sum_i b'_i)} \underbrace{f(s'; s_{\text{nom}}, \sigma_s)}_{\text{constraint signal}} \prod_i \underbrace{f(b'_i; b_i, \sigma_i)}_{\text{constraint background i}}$$

- McLimit/CLsGenerator toss pseudo-experiments in which
  - N_{obs} is sampled from marginal likelihood (hence the hybrid nature of the method)

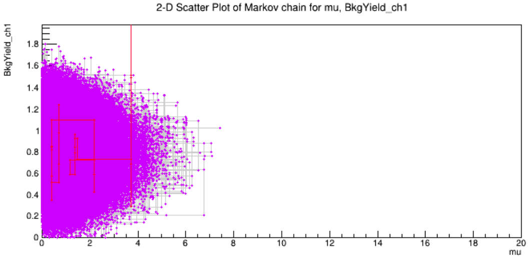
$$\mathscr{L}(\mu) = \int \frac{(\mu s' + \sum_i b'_i)^{N_{\text{obs}}}}{N_{\text{obs}}!} e^{-(\mu s' + \sum_i b'_i)} f(s'; s_{\text{nom}}, \sigma_s) \prod_i f(b'_i; b_i, \sigma_i) \mathrm{d}s' \prod_i \mathrm{d}b'_i \quad (\square)$$

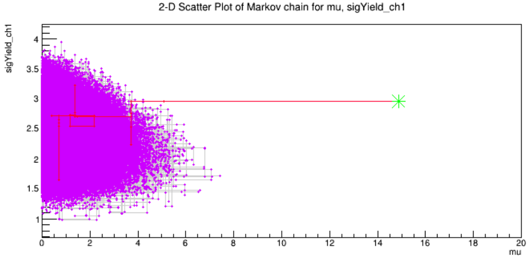
  - Test is computed using nominal values of nuisance parameters

$$\mathscr{L}(\mu, s' = s_{\text{nom}}, \{b'_i\} = \{b_i\}) = \frac{(\mu s_{\text{nom}} + \sum_i b_i)^{N_{\text{obs}}}}{N_{\text{obs}}!} e^{-(\mu s_{\text{nom}} + \sum_i b_i)}$$

- BayesianMCMC uses marginal likelihood ($\square$) as posterior (integration done by Markov Chain Monte Carlo)

Visualization of Markov Chain



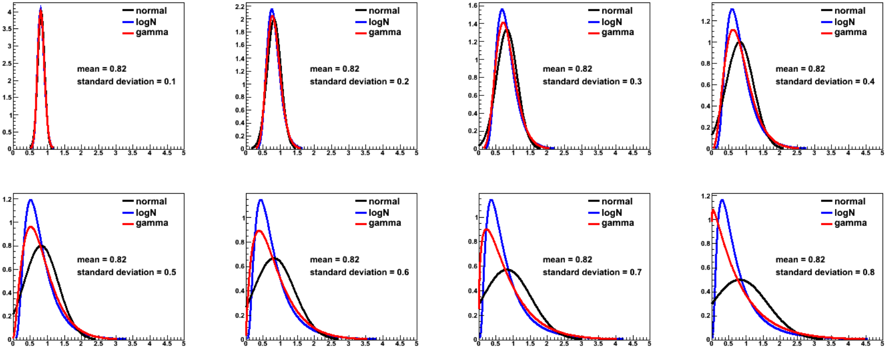


Choice of constraint terms

- This choice is to some extend arbitrary
- However, care must be taken because some choices behave sometimes badly
- Common choices

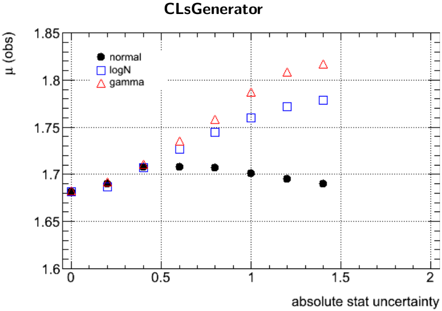
Constraint p. d. f.	**Available in**
Normal	McLimit, CLsGenerator, BayesianMCMC
Log-normal	CLsGenerator, BayesianMCMC
Gamma	CLsGenerator, BayesianMCMC

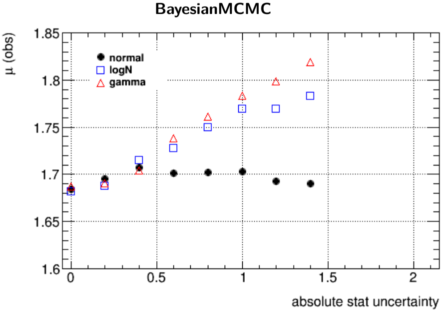
Comparison of distributions



Effect of statistical uncertainty and constraint p. d. f.

- $N_{\text{obs}} = 1,\ b = 0.82$ and $s_{\text{nom}} = 2.49$
- Statistical uncertainty on $b = 0, \ldots, 1.4$





- CLsGenerator and BayesianMCMC give identical results
  - It can be shown that it's always the case when stat. uncert. on signal=0 !

Proof of equivalence between CL_s and Bayesian

- CL_s method

$$\alpha = CL_s(\mu_{\rm up}) = \frac{CL_{s+b}(\mu_{\rm up})}{CL_b} = \frac{\sum\limits_{N=0}^{N_{\rm obs}} \mathscr{L}(\mu_{\rm up})}{\sum\limits_{N=0}^{N_{\rm obs}} \mathscr{L}(\mu{=}0)} =$$

$$\frac{\displaystyle\int \sum_{N=0}^{N_{\rm obs}} \frac{\left(\mu_{\rm up} s' + \sum_i b_i'\right)^N}{N!} e^{-(\mu_{\rm up} s' + \sum_i b_i')} f(s'; s_{\rm nom}, \sigma_s) \prod_i f(b_i'; b_i, \sigma_i) \mathrm{d}s' \prod_i \mathrm{d}b_i'}{\displaystyle\int \sum_{N=0}^{N_{\rm obs}} \frac{\left(\sum b_i'\right)^N}{N!} e^{-\sum_i b_i'} f(s'; s_{\rm nom}, \sigma_s) \prod_i f(b_i'; b_i, \sigma_i) \mathrm{d}s' \prod_i \mathrm{d}b_i'} =$$

$$\frac{\displaystyle\int CL_{s+b}(\mu_{\rm up}, s', \{b_i'\}) f(s'; s_{\rm nom}, \sigma_s) \prod_i f(b_i'; b_i, \sigma_i) \mathrm{d}s' \prod_i \mathrm{d}b_i'}{\displaystyle\int CL_b(\{b_i'\}) f(s'; s_{\rm nom}, \sigma_s) \prod_i f(b_i'; b_i, \sigma_i) \mathrm{d}s' \prod_i \mathrm{d}b_i'}$$

- Bayesian w/ uniform prior

$$1-\alpha = \frac{\displaystyle\int_0^{\mu_{\rm up}} \mathscr{L}(\mu) \mathrm{d}\mu}{\displaystyle\int_0^{\infty} \mathscr{L}(\mu) \mathrm{d}\mu} = \frac{\displaystyle\int \left[\int_0^{\mu_{\rm up}} \frac{\left(\mu s' + \sum_i b_i'\right)^{N_{\rm obs}}}{N_{\rm obs}!} e^{-(\mu s' + \sum_i b_i')} \mathrm{d}\mu \right] f(s'; s_{\rm nom}, \sigma_s) \prod_i f(b_i'; b_i, \sigma_i) \mathrm{d}s' \prod_i \mathrm{d}b_i'}{\displaystyle\int \left[\int_0^{\infty} \frac{\left(\mu s' + \sum_i b_i'\right)^{N_{\rm obs}}}{N_{\rm obs}!} e^{-(\mu s' + \sum_i b_i')} \mathrm{d}\mu \right] f(s'; s_{\rm nom}, \sigma_s) \prod_i f(b_i'; b_i, \sigma_i) \mathrm{d}s' \prod_i \mathrm{d}b_i'} =$$

$$\frac{\displaystyle\int \left[\frac{\Gamma(N_{\rm obs}+1; \sum b_i') - \Gamma(N_{\rm obs}+1; \mu_{\rm up} s' + \sum b_i')}{s' \Gamma(N_{\rm obs}+1)} \right] f(s'; s_{\rm nom}, \sigma_s) \prod_i f(b_i'; b_i, \sigma_i) \mathrm{d}s' \prod_i \mathrm{d}b_i'}{\displaystyle\int \left[\frac{\Gamma(N_{\rm obs}+1; \sum b_i')}{s' \Gamma(N_{\rm obs}+1)} \right] f(s'; s_{\rm nom}, \sigma_s) \prod_i f(b_i'; b_i, \sigma_i) \mathrm{d}s' \prod_i \mathrm{d}b_i'}$$

Proof of equivalence between CL_s and Bayesian

- Bayesian w/ uniform prior (continued)

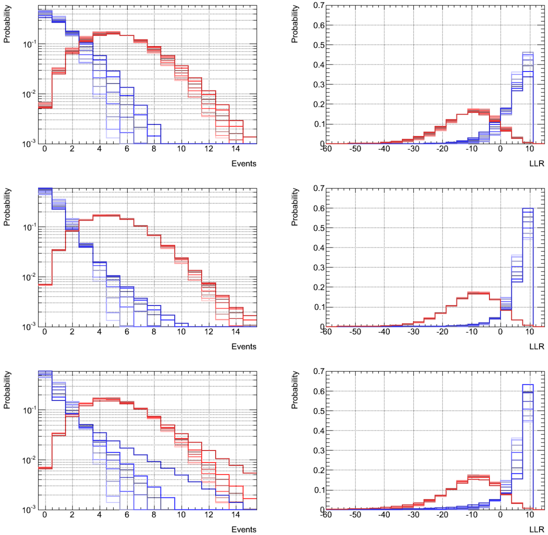
$$\Rightarrow 1-\alpha = 1 - \frac{\int \left[\frac{\Gamma(N_{\mathrm{obs}+1};\mu_{\mathrm{up}} s' + \sum b'_l)}{s'\Gamma(N_{\mathrm{obs}}+1)} \right] f(s';s_{\mathrm{nom}},\sigma_s) \prod_i f(b'_i;b_i,\sigma_i) \mathrm{d}s' \prod_i \mathrm{d}b'_i}{\int \left[\frac{\Gamma(N_{\mathrm{obs}+1};\sum b'_l)}{s'\Gamma(N_{\mathrm{obs}}+1)} \right] f(s';s_{\mathrm{nom}},\sigma_s) \prod_i f(b'_i;b_i,\sigma_i) \mathrm{d}s' \prod_i \mathrm{d}b'_i} =$$

$$1 - \frac{\int \left[\frac{CL_{s+b}(\mu_{\mathrm{up}},s',\{b'_i\})}{s'} \right] f(s';s_{\mathrm{nom}},\sigma_s) \prod_i f(b'_i;b_i,\sigma_i) \mathrm{d}s' \prod_i \mathrm{d}b'_i}{\int \left[\frac{CL_b(\{b'_i\})}{s'} \right] f(s';s_{\mathrm{nom}},\sigma_s) \prod_i f(b'_i;b_i,\sigma_i) \mathrm{d}s' \prod_i \mathrm{d}b'_i}$$

- CL_s and bayesian w/ uniform prior are equivalent if signal perfectly known: $s' = s_{\mathrm{nom}}$

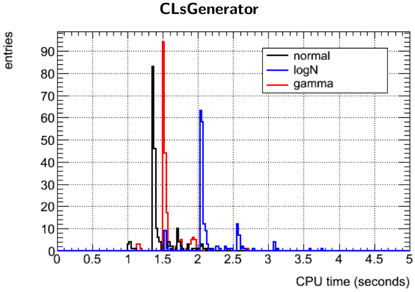
Effect of statistical uncertainty and constraint p. d. f.

normal (top), logN (middle) and gamma (bottom)



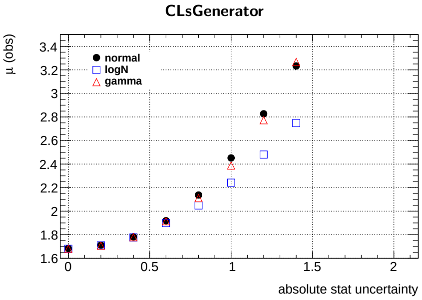
Effect of statistical uncertainty and constraint p. d. f.

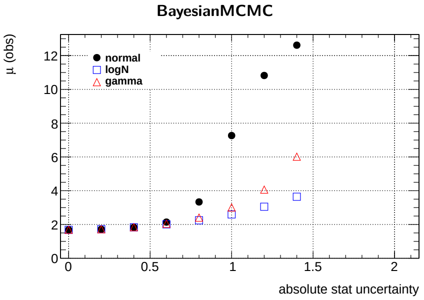
- Truncation at 0 in normal case doesn't smear distributions but shifts them
  - Happens for both b and s+b distribs $\Rightarrow \mu$ doesn't change much
- logN smears less than gamma $\Rightarrow$ better μ for logN than for gamma
- What is the best choice ? logN ?



Effect of statistical uncertainty and constraint p. d. f.

- $N_{\text{obs}} = 1,\ b = 0.82$ and $s_{\text{nom}} = 2.49$
- Statistical uncertainty on s_{nom} and $b = 0, \ldots, 1.4$

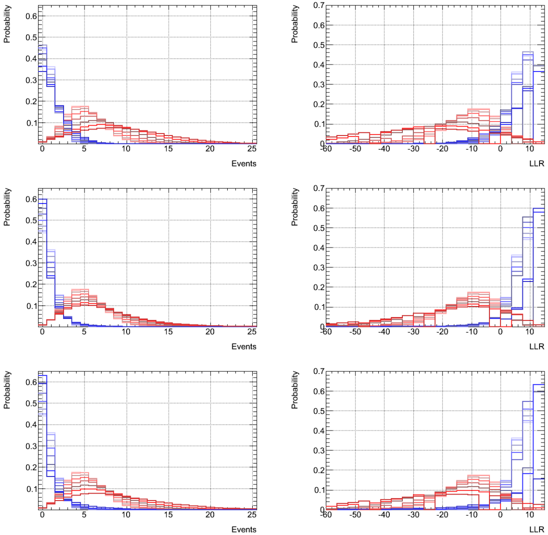




- CLsGenerator and BayesianMCMC give very different results for large uncertainties

Effect of statistical uncertainty and constraint p. d. f.

normal (top), logN (middle) and gamma (bottom)



Effect of statistical uncertainty and constraint p. d. f.

- logN smears less than gamma $\Rightarrow$ better μ for logN than for gamma (as in previous case)
- Normal case now very different from previous case
  - s+b distribution is now "more smeared" than b alone distribution $\Rightarrow$ have to go to larger μ to reach given confidence level

Systematic uncertainties

- Systematic uncertainties more complicated than statistical ones
  - Correlations between samples and channels
  - Interpolation/extrapolation problem

Interpolation/extrapolation of systematics

- For one sample and one systematic we know
  - Nominal yield: N_{nom}
  - Yield with systematic varied 1σ up: $N_{\uparrow}$
  - Yield with systematic varied 1σ down: $N_{\downarrow}$

- But we need a continuous parametrization: $N(\eta)$
  - η defined such that
    - $N(\eta = 0) = N_{\text{nom}}$
    - $N(\eta = +1) = N_{\uparrow}$
    - $N(\eta = -1) = N_{\downarrow}$

- How to interpolate and extrapolate ?

Interpolation/extrapolation of systematics

- For one sample and one systematic we know
  - Nominal yield: N_{nom}
  - Yield with systematic varied 1σ up: $N_{\uparrow}$
  - Yield with systematic varied 1σ down: $N_{\downarrow}$

- But we need a continuous parametrization: $N(\eta)$
  - η defined such that
    - $N(\eta = 0) = N_{\text{nom}}$
    - $N(\eta = +1) = N_{\uparrow}$
    - $N(\eta = -1) = N_{\downarrow}$

- How to interpolate and extrapolate ?

Interpolation/extrapolation of systematics

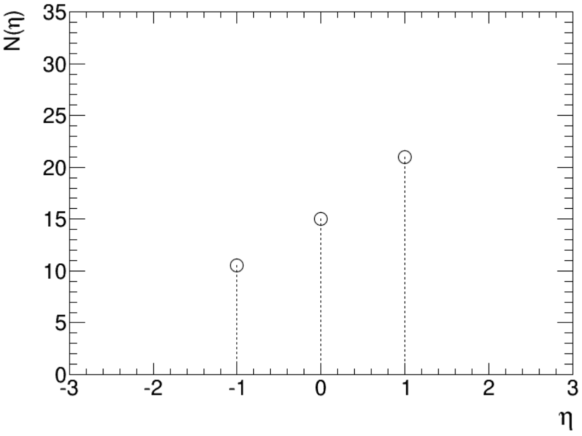
- For one sample and one systematic we know
  - Nominal yield: N_{nom}
  - Yield with systematic varied 1σ up: $N_{\uparrow}$
  - Yield with systematic varied 1σ down: $N_{\downarrow}$

- But we need a continuous parametrization: $N(\eta)$
  - η defined such that
    - $N(\eta = 0) = N_{\text{nom}}$
    - $N(\eta = +1) = N_{\uparrow}$
    - $N(\eta = -1) = N_{\downarrow}$

- How to interpolate and extrapolate ?

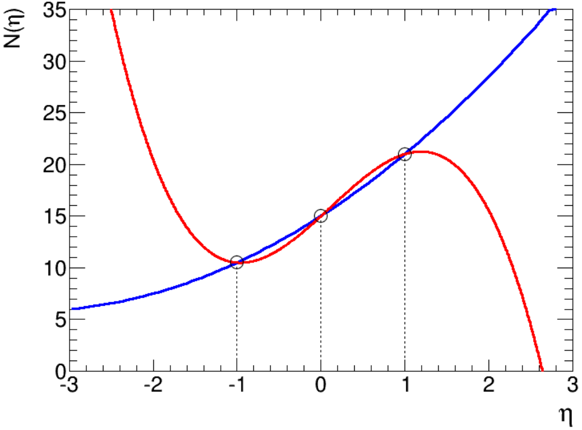
Interpolation/extrapolation of systematics

Example : $N_{\text{nom}} = 15,\ N_{\uparrow} = 21,\ N_{\downarrow} = 10.5$



Interpolation/extrapolation of systematics

Example : $N_{\mathrm{nom}} = 15,\ N_{\uparrow} = 21,\ N_{\downarrow} = 10.5$



Interpolation/extrapolation of systematics

- Rather than using $N_{\text{nom},\uparrow,\downarrow}$, let's use
  - $f^{\uparrow} = \frac{N_{\uparrow} - N_{\text{nom}}}{N_{\text{nom}}}$
  - $f^{\downarrow} = \frac{N_{\downarrow} - N_{\text{nom}}}{N_{\text{nom}}}$
  - $f^{\text{syst}}(\eta) = \frac{N(\eta)}{N_{\text{nom}}}$
- One has
  - $f^{\text{syst}}(\eta = 0) = 1$
  - $f^{\text{syst}}(\eta = -1) = 1 + f^{\downarrow}$
  - $f^{\text{syst}}(\eta = 1) = 1 + f^{\uparrow}$
- Goal: find an inter/extrapolation algorithm such that these relations are satisfied (at least approximately)

Interpolation/extrapolation of systematics

- Rather than using $N_{\text{nom},\uparrow,\downarrow}$, let's use
  - $f^{\uparrow} = \frac{N_{\uparrow} - N_{\text{nom}}}{N_{\text{nom}}}$
  - $f^{\downarrow} = \frac{N_{\downarrow} - N_{\text{nom}}}{N_{\text{nom}}}$
  - $f^{\text{syst}}(\eta) = \frac{N(\eta)}{N_{\text{nom}}}$
- One has
  - $f^{\text{syst}}(\eta = 0) = 1$
  - $f^{\text{syst}}(\eta = -1) = 1 + f^{\downarrow}$
  - $f^{\text{syst}}(\eta = 1) = 1 + f^{\uparrow}$
- Goal: find an inter/extrapolation algorithm such that these relations are satisfied (at least approximately)

Interpolation/extrapolation of systematics

- Rather than using $N_{\text{nom},\uparrow,\downarrow}$, let's use
  - $f^{\uparrow} = \frac{N_{\uparrow} - N_{\text{nom}}}{N_{\text{nom}}}$
  - $f^{\downarrow} = \frac{N_{\downarrow} - N_{\text{nom}}}{N_{\text{nom}}}$
  - $f^{\text{syst}}(\eta) = \frac{N(\eta)}{N_{\text{nom}}}$
- One has
  - $f^{\text{syst}}(\eta = 0) = 1$
  - $f^{\text{syst}}(\eta = -1) = 1 + f^{\downarrow}$
  - $f^{\text{syst}}(\eta = 1) = 1 + f^{\uparrow}$
- Goal: find an inter/extrapolation algorithm such that these relations are satisfied (at least approximately)

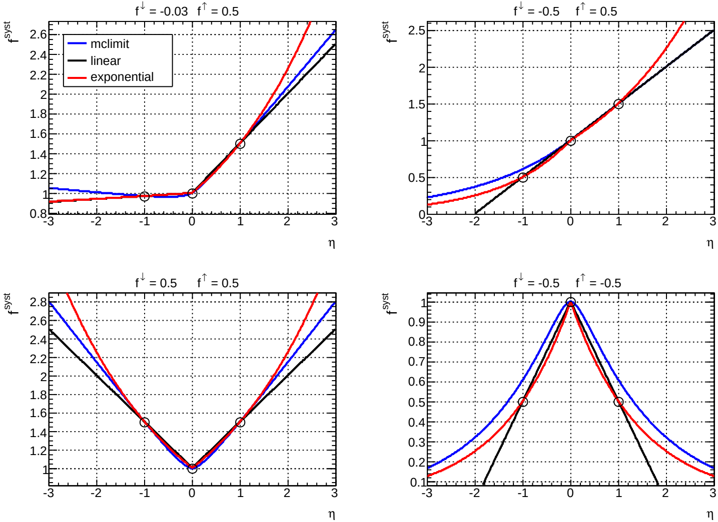
Interpolation/extrapolation of systematics

- Inter/extrapolation algorithms investigated here

Algorithm	Available in
mclimit	McLimit, CLsGenerator
linear	CLsGenerator, BayesianMCMC, RooStats
exponential	CLsGenerator, BayesianMCMC, RooStats

- Remark: other choices available in RooStats and BayesianMCMC
  - RooStats default: polynomial interpolation/exponential extrapolation

First look at mclimit, linear and exponential algorithms



Closer look at mclimit algorithm

- What does mclimit give for $\eta = -1, 0, +1$?
  - $f^{\text{syst}}(\eta = 0) = 1$
  - if $f^{\uparrow} \geq 0$: $f^{\text{syst}}(\eta = +1) = 1 + f^{\uparrow}$
  - if $f^{\downarrow} \geq 0$: $f^{\text{syst}}(\eta = -1) = 1 + f^{\downarrow}$
  - if $f^{\uparrow} < 0$: $f^{\text{syst}}(\eta = +1) = e^{f^{\uparrow}}$
  - if $f^{\downarrow} < 0$: $f^{\text{syst}}(\eta = -1) = e^{f^{\downarrow}}$
- In the last two cases, $f^{\text{syst}} \simeq 1 + f^{\uparrow(\downarrow)}$ only if $f^{\uparrow(\downarrow)} \simeq 0$ (if $f^{\uparrow(\downarrow)} \simeq -1$, difference can be large).

Closer look at mclimit algorithm

- What does mclimit give when uncertainties are symetric: $f^{\uparrow} = -f^{\downarrow}$?
  - $\eta > 0$:
    - $f^{\uparrow} \geq 0$: $f^{syst} = 1 + \eta f^{\uparrow}$
    - $f^{\uparrow} < 0$: $f^{syst} = e^{\eta f^{\uparrow}}$
  - $\eta < 0$:
    - $f^{\downarrow} \geq 0$: $f^{syst} = 1 - \eta f^{\downarrow}$
    - $f^{\downarrow} < 0$: $f^{syst} = e^{-\eta f^{\downarrow}}$
- mclimit is equivalent to linear for $\eta > 0$ if $f^{\uparrow} \geq 0$ or $\eta < 0$ if $f^{\downarrow} \geq 0$

Statistical model

- Consider treatment of syst. uncert. as done in McLimit, CLsGenerator and BayesianMCMC
- For each systematic, introduce a nuisance parameter η_j
- Likelihood

Single channel

$$\boxed{\mathscr{L}(\mu,\{\eta_j\}) = \frac{(\mu s' + \sum_i b_i')^{N_{\text{obs}}}}{N_{\text{obs}}!} e^{-(\mu s' + \sum_i b_i')} \prod_j \underbrace{g(\eta_j)}_{\text{constraint systematic } j}}$$

where

- $b_i' = b_i'(\{\eta_j\}) = b_i \times \prod_j f_{ij}^{\text{syst}}(\eta_j)$
- $s' = s'(\{\eta_j\}) = s_{\text{nom}} \times \prod_j f_j^{\text{syst}}(\eta_j)$
- $g(\eta_j) = \mathscr{N}(0,1)$ (standard normal distribution)

Statistical model

- Consider treatment of syst. uncert. as done in McLimit, CLsGenerator and BayesianMCMC
- For each systematic, introduce a nuisance parameter η_j
- Likelihood

Single channel

$$\boxed{\mathscr{L}(\mu,\{\eta_j\}) = \frac{(\mu s' + \sum_i b'_i)^{N_{\text{obs}}}}{N_{\text{obs}}!} e^{-(\mu s' + \sum_i b'_i)} \prod_j \underbrace{g(\eta_j)}_{\text{constraint systematic } j}}$$

where

- $b'_i = b'_i(\{\eta_j\}) = b_i \times \prod_j f_{ij}^{\text{syst}}(\eta_j)$
- $s' = s'(\{\eta_j\}) = s_{\text{nom}} \times \prod_j f_j^{\text{syst}}(\eta_j)$
- $g(\eta_j) = \mathcal{N}(0,1)$ (standard normal distribution)

Statistical model

- Consider treatment of syst. uncert. as done in McLimit, CLsGenerator and BayesianMCMC
- For each systematic, introduce a nuisance parameter η_j
- Likelihood

Single channel

$$\boxed{\mathscr{L}(\mu,\{\eta_j\}) = \frac{(\mu s' + \sum_i b_i')^{N_{\text{obs}}}}{N_{\text{obs}}!} e^{-(\mu s' + \sum_i b_i')} \prod_j \underbrace{g(\eta_j)}_{\text{constraint systematic } j}}$$

where

- $b_i' = b_i'(\{\eta_j\}) = b_i \times \prod_j f_{ij}^{\text{syst}}(\eta_j)$
- $s' = s'(\{\eta_j\}) = s_{\text{nom}} \times \prod_j f_j^{\text{syst}}(\eta_j)$
- $g(\eta_j) = \mathscr{N}(0,1)$ (standard normal distribution)

What McLimit/CLsGenerator/BayesianMCMC do ?

Single channel

$$\boxed{\mathscr{L}(\mu,\{\eta_j\}) = \frac{(\mu s' + \sum_i b_i')^{N_{\rm obs}}}{N_{\rm obs}!} e^{-(\mu s' + \sum_i b_i')} \prod_j \underbrace{g(\eta_j)}_{\text{constraint systematic } j}}$$

- McLimit/CLsGenerator toss pseudo-experiments in which
  - $N_{\rm obs}$ is sampled from marginal likelihood (hence the hybrid nature of the method)

$$\mathscr{L}(\mu) = \int \frac{(\mu s' + \sum_i b_i')^{N_{\rm obs}}}{N_{\rm obs}!} e^{-(\mu s' + \sum_i b_i')} \prod_j g(\eta_j) d\eta_j \quad (\triangle)$$

  - Test is computed using nominal values of nuisance parameters

$$\mathscr{L}(\mu,\{\eta_j\} = 0) = \frac{(\mu s_{\rm nom} + \sum_i b_i)^{N_{\rm obs}}}{N_{\rm obs}!} e^{-(\mu s_{\rm nom} + \sum_i b_i)}$$

- BayesianMCMC uses marginal likelihood ($\triangle$) as posterior (integration done by Markov Chain Monte Carlo)

What McLimit/CLsGenerator/BayesianMCMC do ?

Single channel

$$\boxed{\mathscr{L}(\mu,\{\eta_j\}) = \frac{(\mu s' + \sum_i b_i')^{N_{\text{obs}}}}{N_{\text{obs}}!} e^{-(\mu s' + \sum_i b_i')} \prod_j \underbrace{g(\eta_j)}_{\text{constraint systematic } j}}$$

- McLimit/CLsGenerator toss pseudo-experiments in which
  - N_{obs} is sampled from marginal likelihood (hence the hybrid nature of the method)

$$\mathscr{L}(\mu) = \int \frac{(\mu s' + \sum_i b_i')^{N_{\text{obs}}}}{N_{\text{obs}}!} e^{-(\mu s' + \sum_i b_i')} \prod_j \mathrm{g}(\eta_j)\mathrm{d}\eta_j \quad (\triangle)$$

  - Test is computed using nominal values of nuisance parameters

$$\mathscr{L}(\mu,\{\eta_j\} = 0) = \frac{(\mu s_{\text{nom}} + \sum_i b_i)^{N_{\text{obs}}}}{N_{\text{obs}}!} e^{-(\mu s_{\text{nom}} + \sum_i b_i)}$$

- BayesianMCMC uses marginal likelihood ($\triangle$) as posterior (integration done by Markov Chain Monte Carlo)

What McLimit/CLsGenerator/BayesianMCMC do ?

Single channel

$$\boxed{\mathscr{L}(\mu,\{\eta_j\}) = \frac{(\mu s' + \sum_i b'_i)^{N_{\rm obs}}}{N_{\rm obs}!} e^{-(\mu s' + \sum_i b'_i)} \prod_j \underbrace{g(\eta_j)}_{\text{constraint systematic } j}}$$

- McLimit/CLsGenerator toss pseudo-experiments in which
  - $N_{\rm obs}$ is sampled from marginal likelihood (hence the hybrid nature of the method)

$$\mathscr{L}(\mu) = \int \frac{(\mu s' + \sum_i b'_i)^{N_{\rm obs}}}{N_{\rm obs}!} e^{-(\mu s' + \sum_i b'_i)} \prod_j \mathrm{g}(\eta_j) \mathrm{d}\eta_j \quad (\triangle)$$

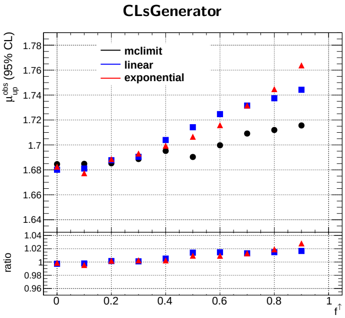
  - Test is computed using nominal values of nuisance parameters

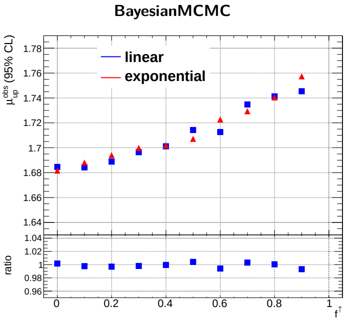
$$\mathscr{L}(\mu,\{\eta_j\} = 0) = \frac{(\mu s_{\rm nom} + \sum_i b_i)^{N_{\rm obs}}}{N_{\rm obs}!} e^{-(\mu s_{\rm nom} + \sum_i b_i)}$$

- BayesianMCMC uses marginal likelihood ($\triangle$) as posterior (integration done by Markov Chain Monte Carlo)

Effect of inter/extrapolation algorithms

- $N_{\mathrm{obs}} = 1,\ b = 0.82$ and $s_{\mathrm{nom}} = 2.49$
- Systematic uncertainty on b: $f^{\uparrow} = -f^{\downarrow} = 0, 0.1, 0.2, \ldots, 0.9$

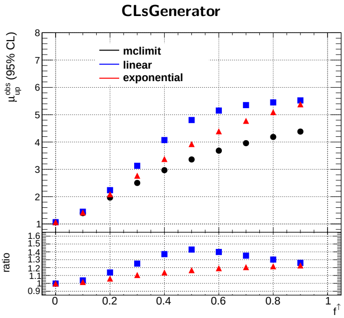


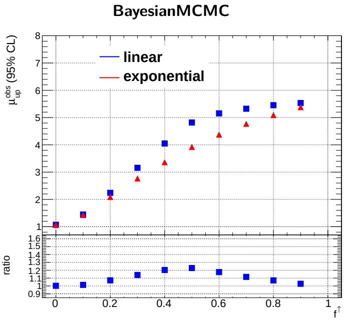


- CLsGenerator and BayesianMCMC give identical results ! (as for stat. uncert.)

Effect of inter/extrapolation algorithms

- $N_{\text{obs}} = 100$, $b = 100$ and $s_{\text{nom}} = 20$
- Systematic uncertainty on b: $f^{\uparrow} = -f^{\downarrow} = 0, 0.1, 0.2, \ldots, 0.9$

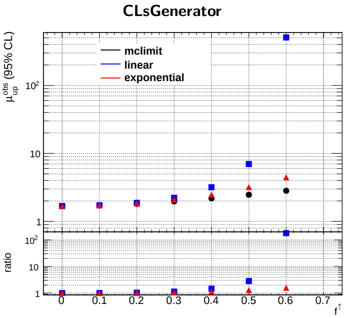


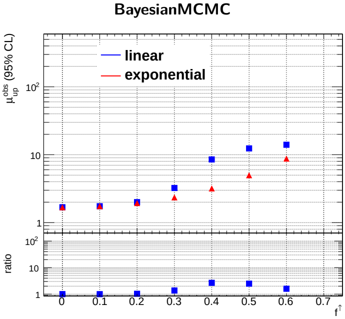


- CLsGenerator and BayesianMCMC give identical results ! (as for stat. uncert.)

Effect of inter/extrapolation algorithms

- $N_{\rm obs} = 1,\ b = 0.82$ and $s_{\rm nom} = 2.49$
- Systematic uncertainty on $s_{\rm nom}$ and b (100% correlated): $f^{\uparrow} = -f^{\downarrow} = 0, \ldots, 0.6$





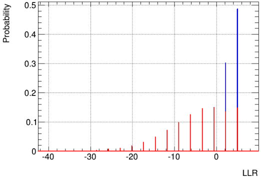
- CLsGenerator and BayesianMCMC give very different results for large systematics

Effect of inter/extrapolation algorithms

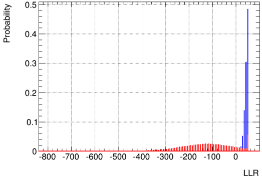
- Comment on CLsGenerator:
  - $q_\mu = 2\left(\mu s_{\text{nom}} - N_{obs} \ln \frac{\mu s_{\text{nom}} + \sum b_i}{\sum b_i}\right)$
  - When $f^{\downarrow} \ll 0$, $f^{syst} = 0$ quite frequently
    $\Rightarrow s = b = 0 \Rightarrow N_{obs} = 0 \Rightarrow q_\mu = 2\mu s$ (for both hypothesis)$\Rightarrow$CL$_{s+b}$ and CL$_s$ can't go to very low values as μ increases

Effect of inter/extrapolation algorithms

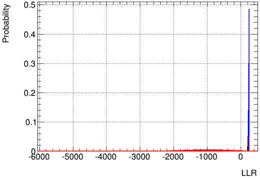
- Consider the point $f^{\uparrow} = -f^{\downarrow} = 0.6$



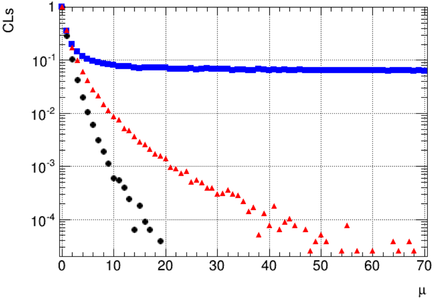
$\mu = 1$



$\mu = 10$



$\mu = 50$



More realistic case: LHCP $\mu\mu$ channel

- Including all stat. and syst. uncertainties

	mclimit	linear	expo
Expected median	1.70	1.70	1.69
Expected $\pm 1\sigma$	1.44-2.08	1.43-2.07	1.43-2.07
Expected $\pm 2\sigma$	1.26-3.52	1.27-3.53	1.27-3.54
Observed	1.67	1.68	1.67

$\rightarrow$ Not much difference

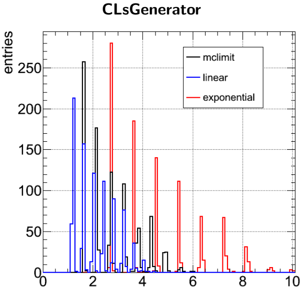
- CPU time

  ran

  clgen.observedSigStrengthFor95excl(0,1e5,cls)

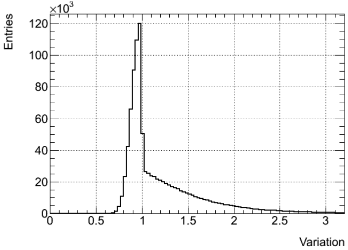
  in identical conditions several

  times to make this distrib



Understanding f^{syst} distribution: exponential case

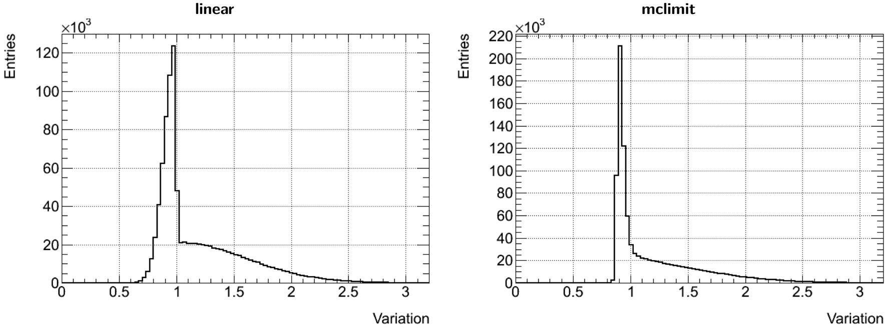
- f^{syst} sometimes looks very weird. Let's consider exponential algo with $f^{\downarrow} = -0.1$ and $f^{\uparrow} = 0.6$ (on the plot, Variation=f^{syst})



- Where does this funny looking shape comes from ?
  - We have $f^{\text{syst}} = (1+f^{\uparrow})^{\eta}$ for $\eta > 0$ ($(1+f^{\uparrow})^{-\eta}$ for $\eta < 0$) with $\eta \sim \mathcal{N}(0,1)$
  - Thus $f^{\text{syst}} \sim \frac{1}{\sqrt{2\pi}|\ln(1+f^{\uparrow(\downarrow)})|f^{syst}} e^{-\frac{1}{2}\left(\frac{\ln f^{syst}}{\ln(1+f^{\uparrow(\downarrow)})}\right)^2}$, i.e. $f^{\text{syst}} \sim$ piecewise $log\mathcal{N}$
- Is such a shape desirable ?

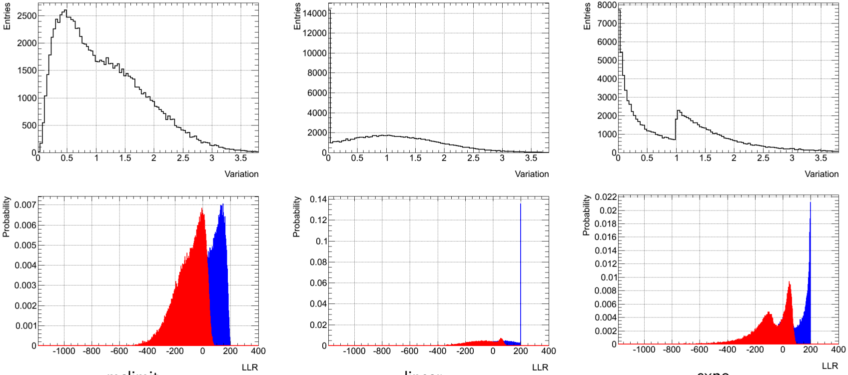
Understanding f^{syst} distribution: linear and mclimit cases

- linear case is obvious: $f^{\text{syst}} \sim$ piecewise gaussian truncated at 0. Using values of previous slide we have
- mclimit case: doesn't seem to be possible to determine analytic solution. Using values of previous slide we have



Extreme cases

- Consider the point $f^{\uparrow} = -f^{\downarrow} = 0.9$ and $\mu = 5.5$ with $N_{\text{obs}} = 100$, $b = 100$ and $s_{\text{nom}} = 20$.



mclimit linear expo

- Using truncated piecewise gaussian (linear) and piecewise $log\mathcal{N}$ (expo) seems absurd $\rightarrow$mclimit much more reasonable in such cases

Summary on inter/extrapolation

- mclimit, linear and exponential algorithms give similar results for small systematics
  - Any choice seems rather safe in this case
- Differences can be significant for large systematics
  - mclimit seems to be the “most reasonable” choice in this case

Summary on inter/extrapolation

- mclimit, linear and exponential algorithms give similar results for small systematics
  - Any choice seems rather safe in this case
- Differences can be significant for large systematics
  - mclimit seems to be the “most reasonable” choice in this case

Summary on inter/extrapolation

- mclimit
  - pros: f^{syst} smooth, $f^{syst}(\eta)\neq 0\ \forall\eta$
  - cons: $f^{syst}(\eta=\pm 1)\neq 1+f^{\uparrow(\downarrow)}$ when $f^{\uparrow(\downarrow)}<0$,
- linear
  - pros: simple, fast
  - cons: $f^{syst}(\eta)=0$ in some cases ($\Rightarrow$ problem when signal syst large)
- expo
  - pros: simple, $f^{syst}(\eta)\neq 0\ \forall\eta$
  - cons: f^{syst} can be very discontinuous for large $f^{\uparrow(\downarrow)}$, slow

Summary on inter/extrapolation

- Systematics are such a pain ! It's good that we're not affected too much by them in the same-sign analysis

Summary of CLsGenerator, BayesianMCMC and McLimit

- Full statistical model

Single channel

$$\boxed{\mathscr{L}(\mu, s', \{b_i'\}, \{\eta_j\}) = \frac{(\mu s'' + \sum_i b_i'')^{N_{\text{obs}}}}{N_{\text{obs}}!} e^{-(\mu s'' + \sum_i b_i'')} f(s'; s_{\text{nom}}, \sigma_s) \prod_i f(b_i'; b_i, \sigma_i) \prod_j g(\eta_j)}$$

with

  - $b_i'' = b_i''(\{\eta_j\}, b_i') = b_i' \times \prod_j f_{ij}^{\text{syst}}(\eta_j)$
  - $s'' = s''(\{\eta_j\}, s') = s' \times \prod_j f_j^{\text{syst}}(\eta_j)$

- Marginal model: likelihood integrated over all nuisance parameters
- McLimit, CLsGenerator
  - Test: $q_\mu = -2 \ln \frac{\mathscr{L}(\mu, s' = s_{\text{nom}}, \{b_i'\} = \{b_i\}, \{\eta_j\} = 0)}{\mathscr{L}(\mu = 0)}$
- BayesianMCMC
  - Integrate marginal likelihood

Conclusion

- Limits are arbitrary in many ways
  - Make cross-checks !
- All checks performed to validate CLsGenerator and BayesianMCMC are successful so far
- CLsGenerator and Bayesian w/ uniform prior are equivalent in Single channel - No uncertainties case
- Analytical solutions exists in Single channel - No uncertainties case
- CLsGenerator and Bayesian w/ uniform prior are equivalent in Single channel - No uncertainties case when no uncertainty on signal

Vrac

- Popper
- Construction Neyman
- Vizualisation Markov Chain