2022 ASIA EUROPE PACIFIC SCHOOL OF HIGH-ENERGY PHYSICS

Practical Statistics

For Particle Physicists

Nicolas Berger (LAPP Annecy)

Lecture 2

Lecture Plan

Statistics basic concepts (Today)

- [**Basic ingredients** (PDFs, etc.)]
- **Statistical Modeling** (PDFs for particle physics measurements)
- **Parameter estimation** (maximum likelihood, least-squares, ...)

Computing statistical results (Today)

- **Model testing** (χ^2 tests, hypothesis testing, p-values, ...)
- **Discovery testing**
- **Confidence intervals**
- **Upper limits**

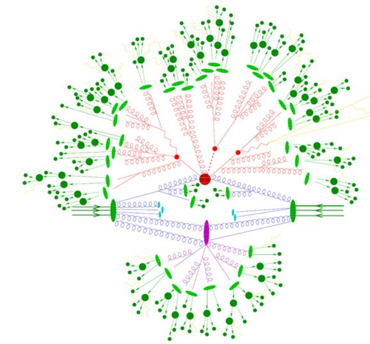
Systematics and further topics (Tomorrow)

- **Systematics and profiling**
- **[Bayesian techniques]**

Disclaimer: the examples and methods covered in the lectures will be biased towards LHC techniques (generally close to the state of the art anyway)

The class will be based on both lectures and hands-on tutorials

Statistical Modeling Reminders



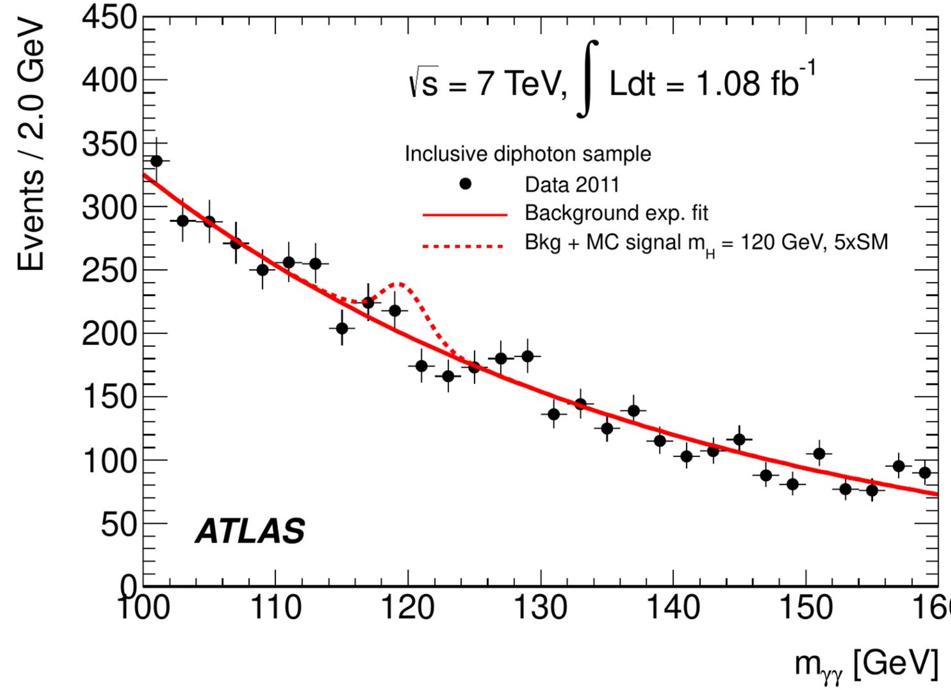
Random data must be described using a statistical model:

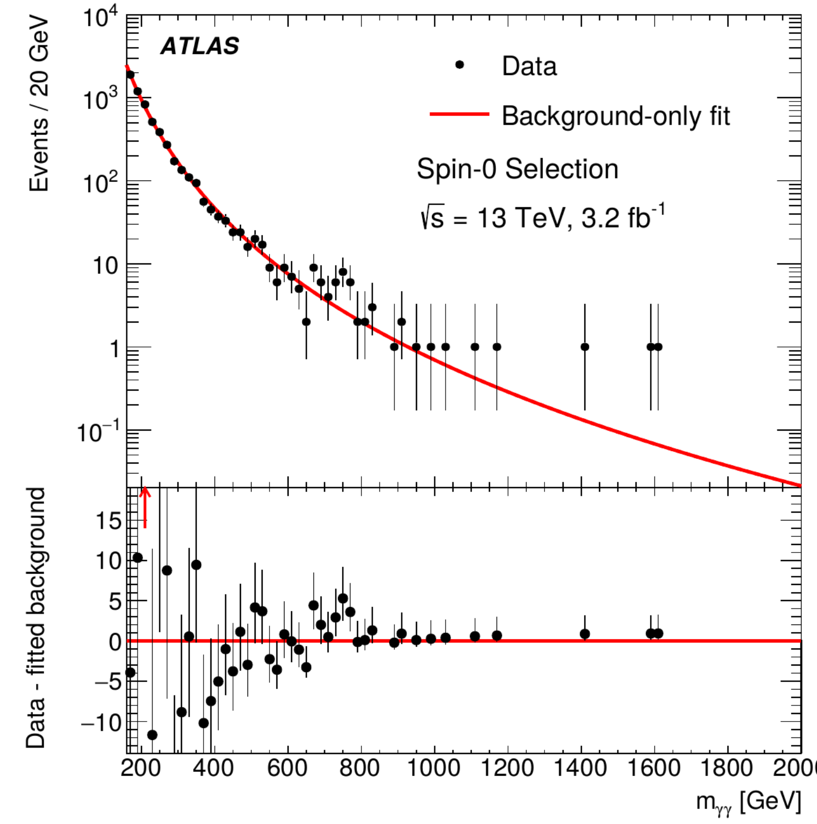
Description	Observable	Likelihood
Counting	n	**Poisson** $$P(n;S,B)=e^{-(S+B)}\frac{(S+B)^{n}}{n!}$$
Binned shape analysis	n_i, i = 1 .. N_{bins}	**Poisson product** $$P(n_i;S,B)=\prod_{i=1}^{n_{\rm bins}} e^{-(S f_i^{\rm sig}+B f_i^{\rm bkg})}\frac{(S f_i^{\rm sig}+B f_i^{\rm bkg})^{n_i}}{n_i!}$$
Unbinned shape analysis	m_i, i = 1 .. n_{evts}	**Extended Unbinned Likelihood** $$P(m_i;S,B)=\frac{e^{-(S+B)}}{n_{\rm evts}!}\prod_{i=1}^{n_{\rm evts}} S\,P_{\rm sig}(m_i)+B\,P_{\rm bkg}(m_i)$$

Includes **parameters of interest** (POIs) but also **nuisance parameters** (NPs)

Next step: use the model to obtain information on the POIs

Hypothesis Testing and discovery

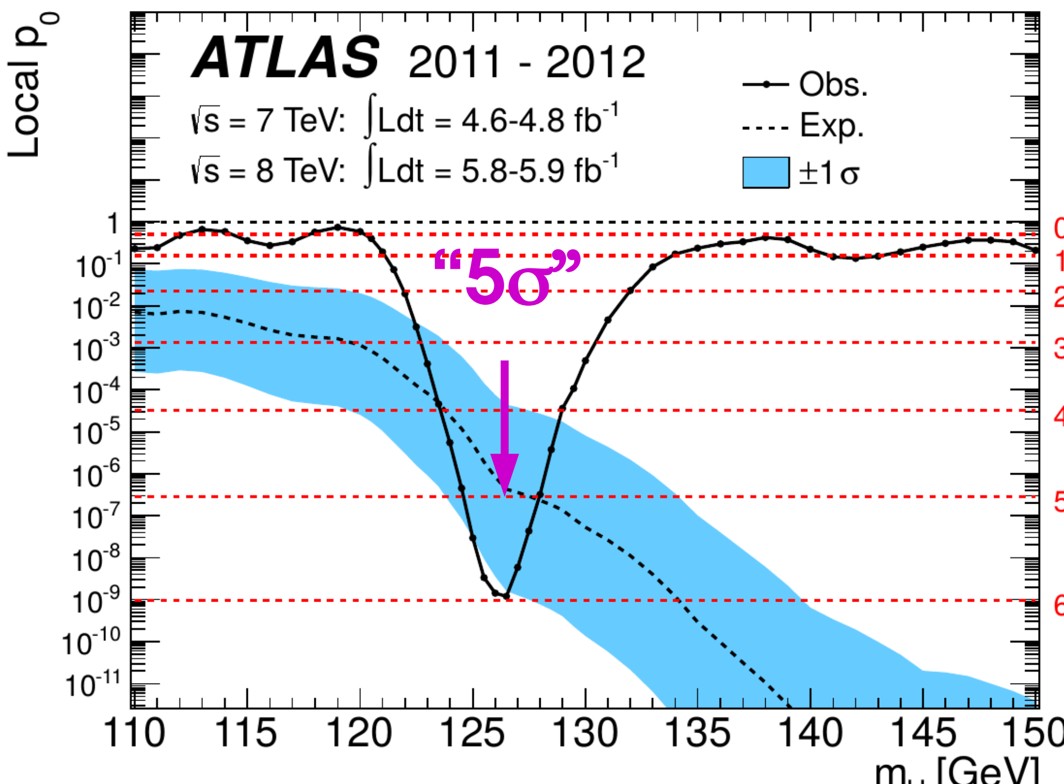




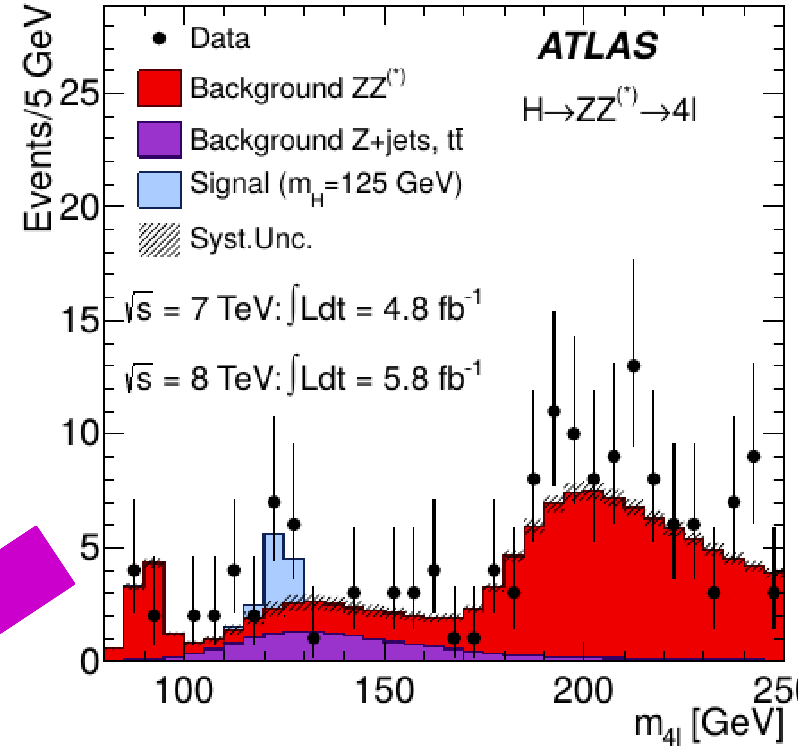
Discovery Testing

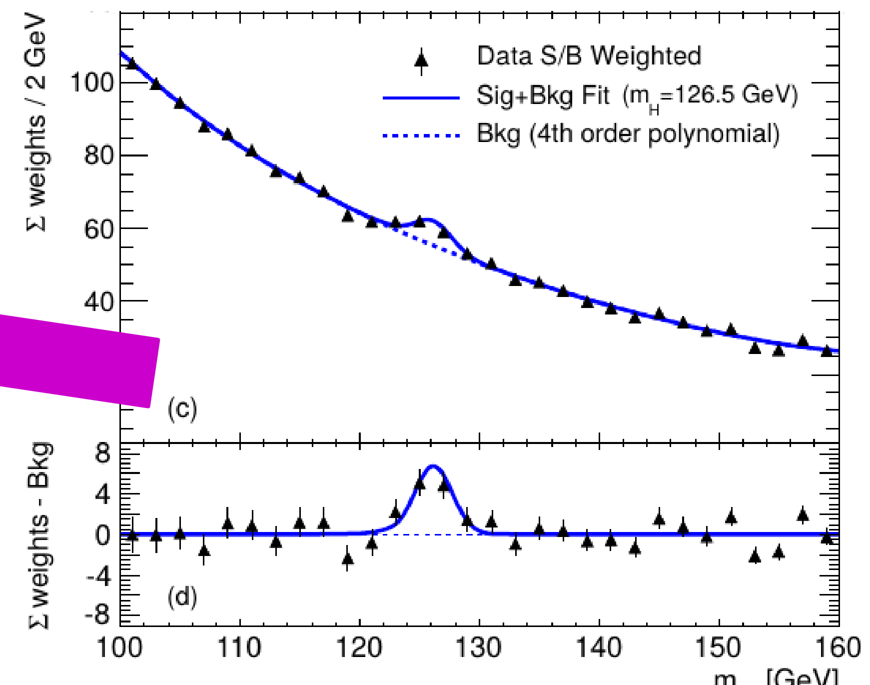
We see an unexpected feature in our data, is it a signal for new physics or a fluctuation ?

e.g. Higgs discovery : **"We have 5σ" !**

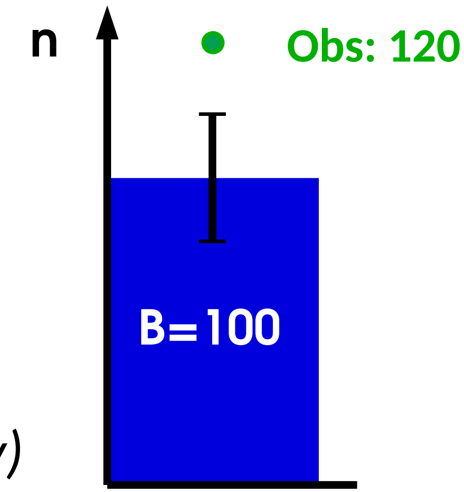


Phys. Lett. B 716 (2012) 1-29





Discovery Testing



Say we have a Gaussian measurement with a background **B=100**, and we measure **n=120**

Did we just discover something ? ***Maybe :-)*** *(but not very likely)*

The measured signal is S = 20.

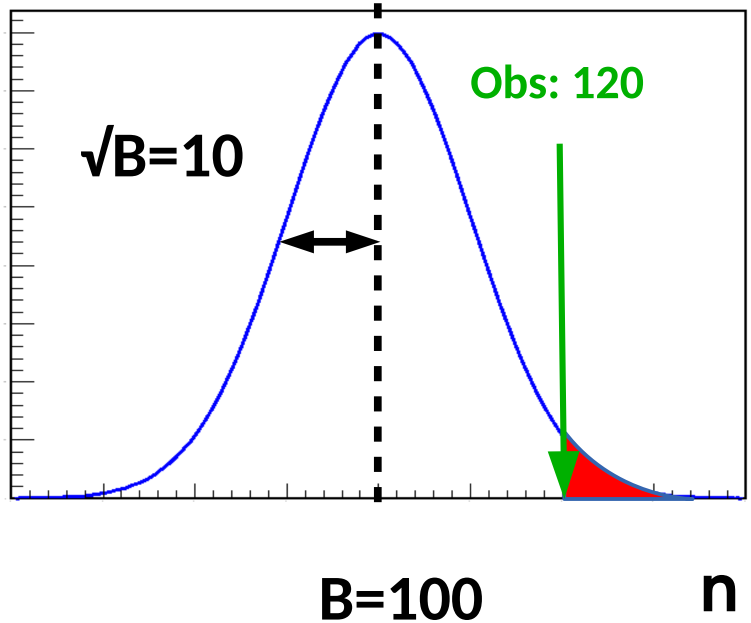
$$S = n_{obs} - B$$

Uncertainty on B is √B = 10

⇒ Significance Z = 2

⇒ we are ~2σ away from S=0.

$$Z = \frac{S}{\sqrt{B}}$$



Gaussian quantiles :

Z = 2 happens p_0 ~ 2.3% of the time if S=0

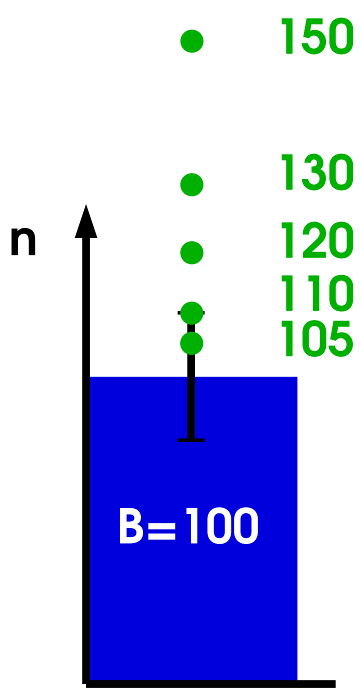
P-value:

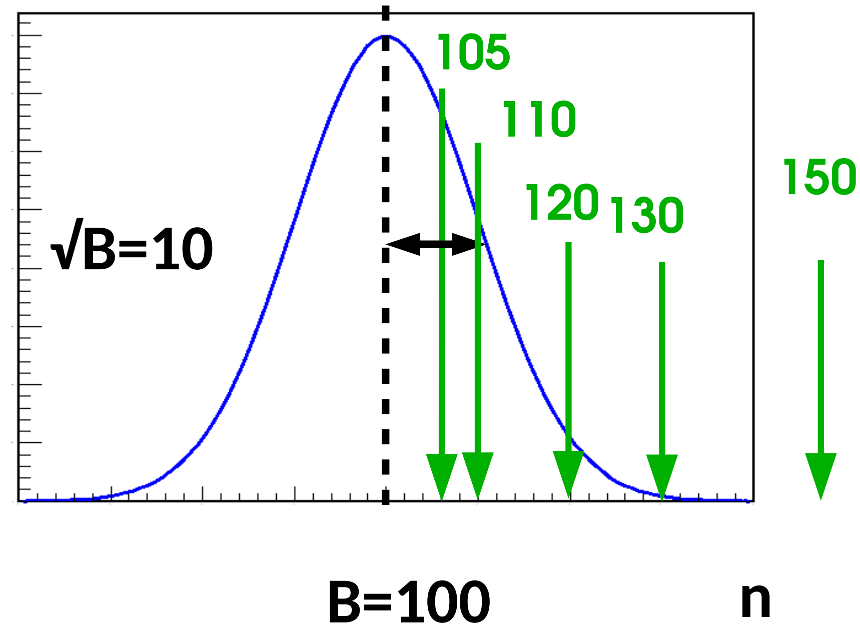
$$p_0 = 1 - \Phi(Z)$$

$$\Phi(Z) = \int_{-\infty}^{Z} G(u;0,1)\, du$$

⇒ Rare, but not exceptional

Discovery Testing





n_{obs}	S	Z	p_0
105	5	0.5σ	31%
110	10	1σ	16%
120	20	2σ	2.3%
130	30	3σ	0.1%
150	50	5σ	$3\ 10^{-7}$

Straightforward in this Gaussian case

Need to be able to do the same in more complex cases:

- **Determine S**
- **Compute Z and p_0**

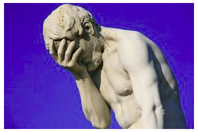
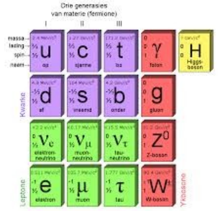
Evidence

Discovery

General Hypothesis Testing

Null Hypothesis: assumption on POIs, say value of S (e.g. **H_0 : S=0**)

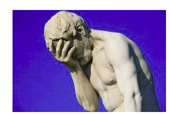
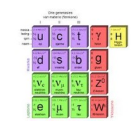
→ **Goal** : decide if H_0 is favored or disfavored using a test based on the data

Possible outcomes:	Data disfavors H_0 (Discovery claim)	Data favors H_0 (Nothing found)
H_0 is false (New physics!)	**Discovery!**  	**Missed discovery** 
H_0 is true (Nothing new)	**False discovery** 	**No new physics, None found** 

"... the null hypothesis is never proved or established, but is possibly disproved, in the course of experimentation. Every experiment may be said to exist only to give the facts a chance of disproving the null hypothesis." – R. A. Fisher

General Hypothesis Testing

Hypothesis: assumption on model parameters, say value of S (e.g. $\mathbf{H_0 : S=0}$)

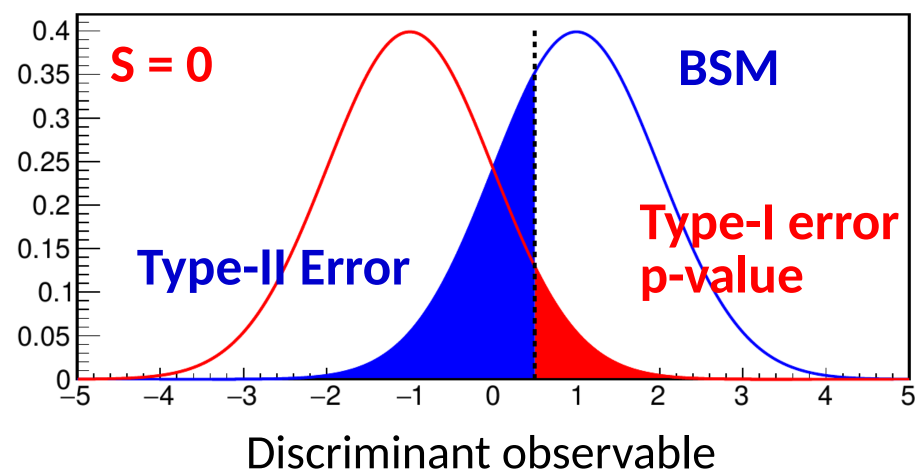
	Data disfavors H_0 (Discovery claim)	Data favors H_0 (Nothing found)
H_0 is false (New physics!)	**Discovery!** 	**Type-II error (Missed discovery)** 
H_0 is true (Nothing new)	**Type-I error (False discovery)** 	**No new physics, none found** 

↳ **p-value, significance**

Lower Type-I errors ⇔ **Higher Type-II errors** and vice versa: cannot have everything!

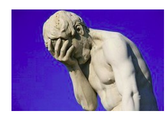
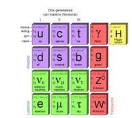
→ **Goal**: test that minimizes Type-II errors **for given level of Type-I error**.

→ Usually set predefined level of **acceptable Type-I error** (e.g. "5σ")



General Hypothesis Testing

Hypothesis: assumption on model parameters, say value of S (e.g. $\mathbf{H_0 : S{=}0}$)

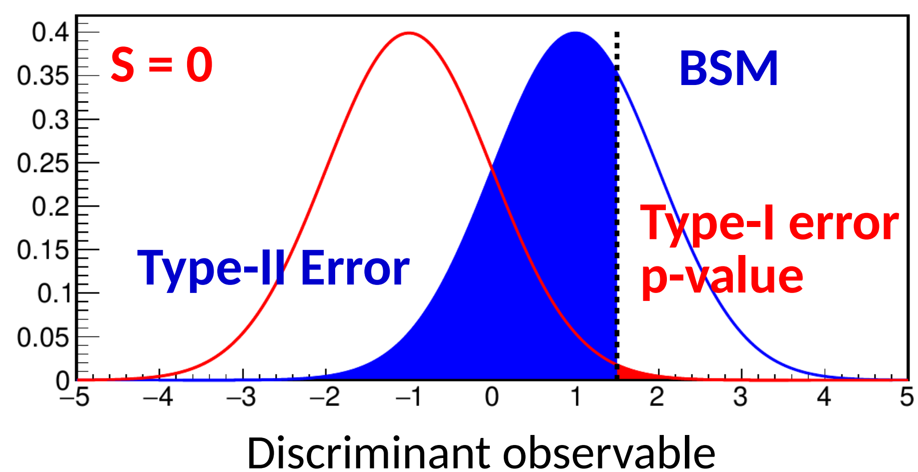
	Data disfavors H_0 (Discovery claim)	Data favors H_0 (Nothing found)
H_0 is false (New physics!)	**Discovery!** 	**Type-II error (Missed discovery)** 
H_0 is true (Nothing new)	**Type-I error (False discovery)** 	**No new physics, none found** 

↑ **p-value, significance**

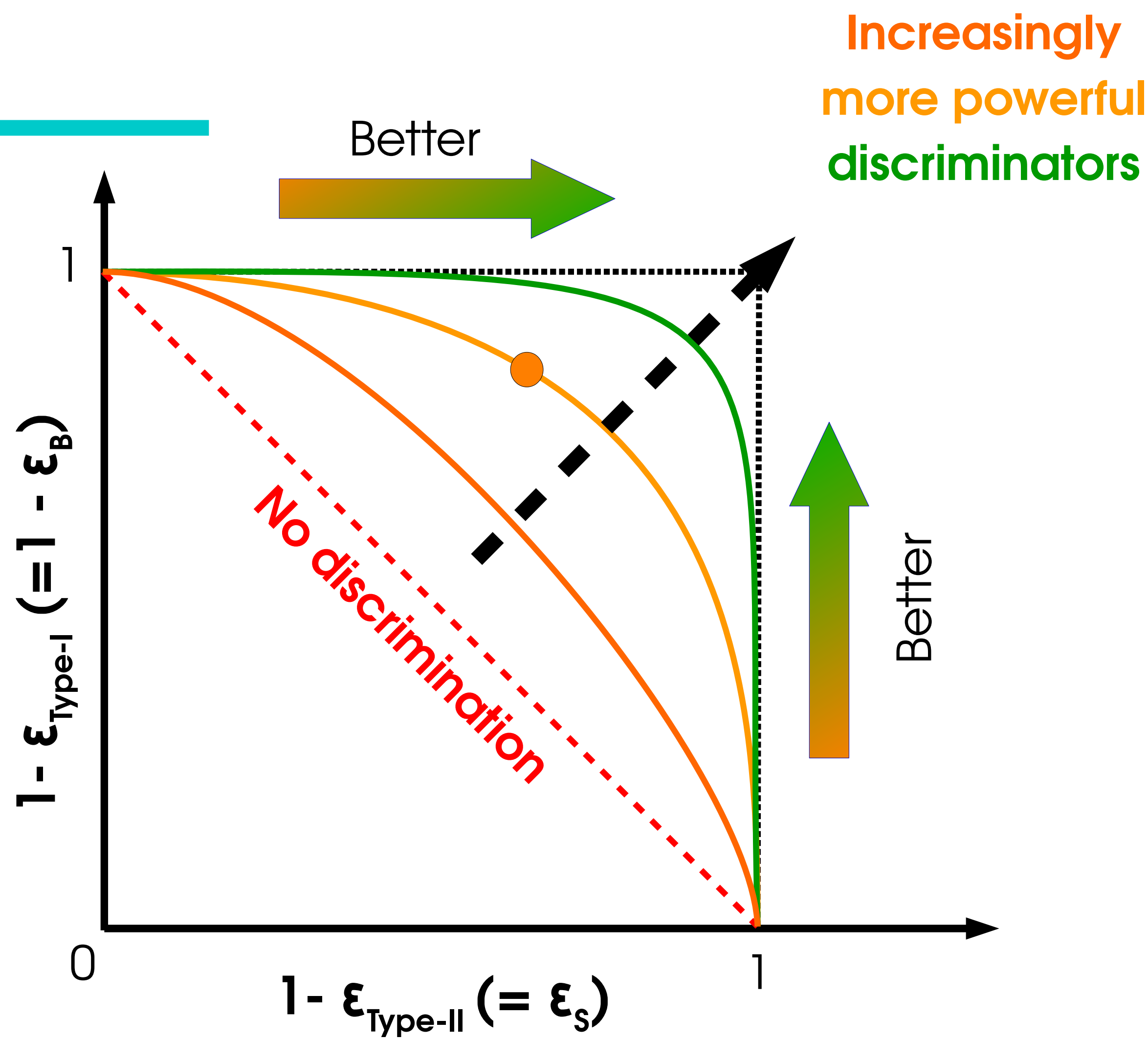
Lower Type-I errors ⇔ **Higher Type-II errors** and vice versa: cannot have everything!

→ **Goal**: test that minimizes Type-II errors **for given level of Type-I error**.

→ Usually set predefined level of **acceptable Type-I error** (e.g. “5σ”)



ROC Curves



"Receiver operating characteristic" (ROC) Curve:

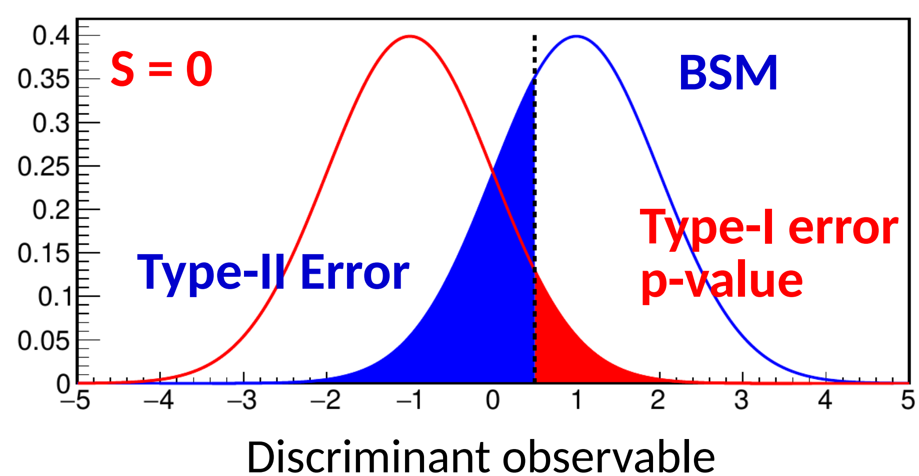
→ Shows Type-I vs Type-II rates for different selections

→ All curves monotonically decrease from (0,1) to (1,0)

→ Better discriminators more bent towards (1,1)

→ **Goal**: test that minimizes Type-II errors **for given level of Type-I error**.

→ Usually set predefined level of **acceptable Type-I error** (e.g. "5σ")



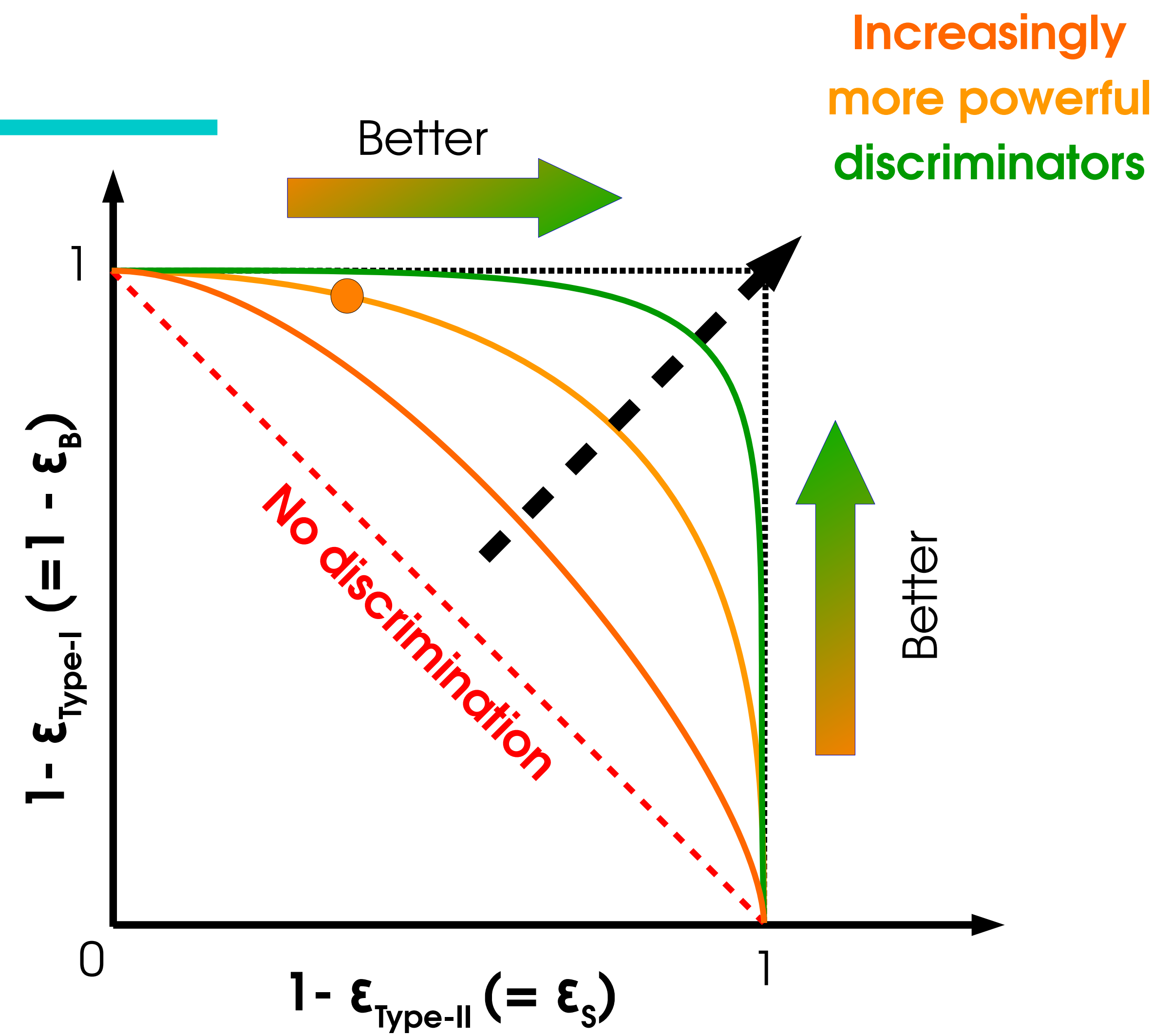
ROC Curves

“Receiver operating characteristic” (ROC) Curve:

→ Shows Type-I vs Type-II rates for different selections

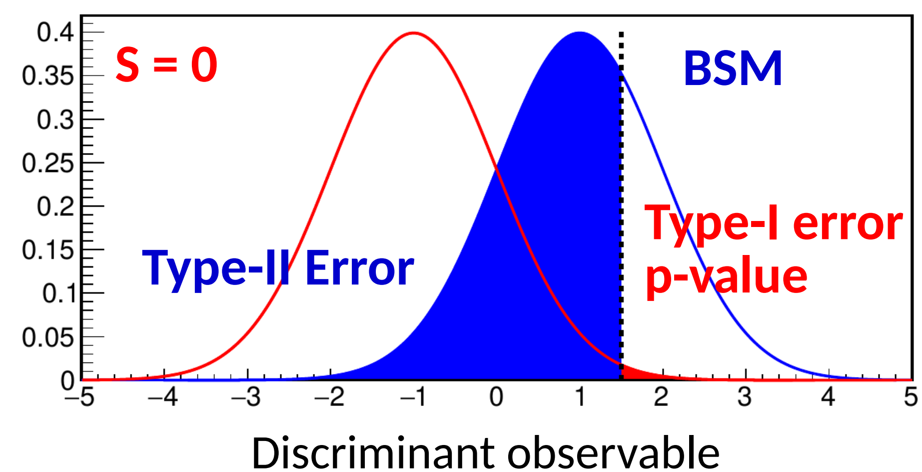
→ All curves monotonically decrease from (0,1) to (1,0)

→ Better discriminators more bent towards (1,1)



→ **Goal**: test that minimizes Type-II errors **for given level of Type-I error**.

→ Usually set predefined level of **acceptable Type-I error** (e.g. “5σ”)



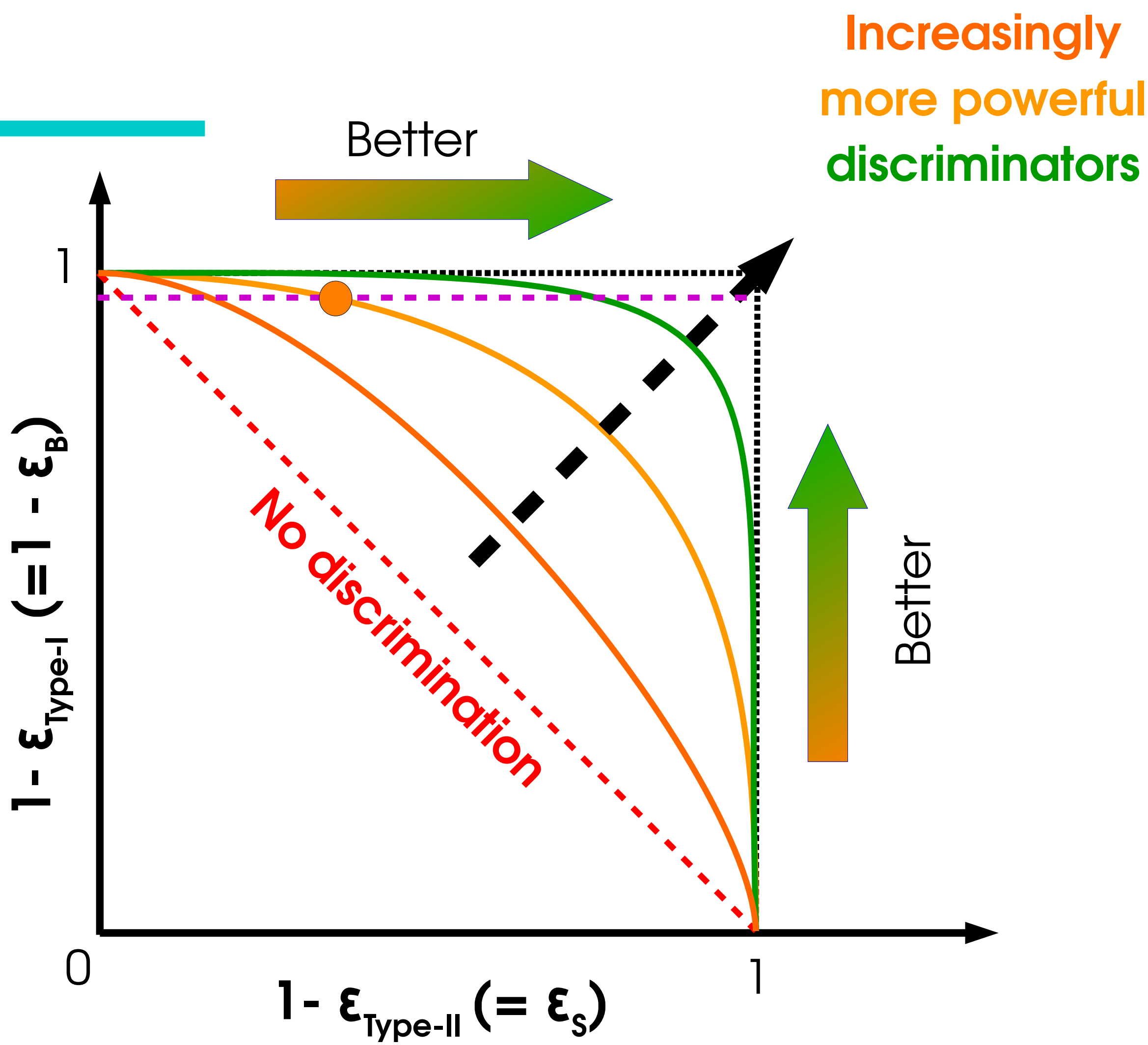
ROC Curves

"Receiver operating characteristic" (ROC) Curve:

→ Shows Type-I vs Type-II rates for different selections

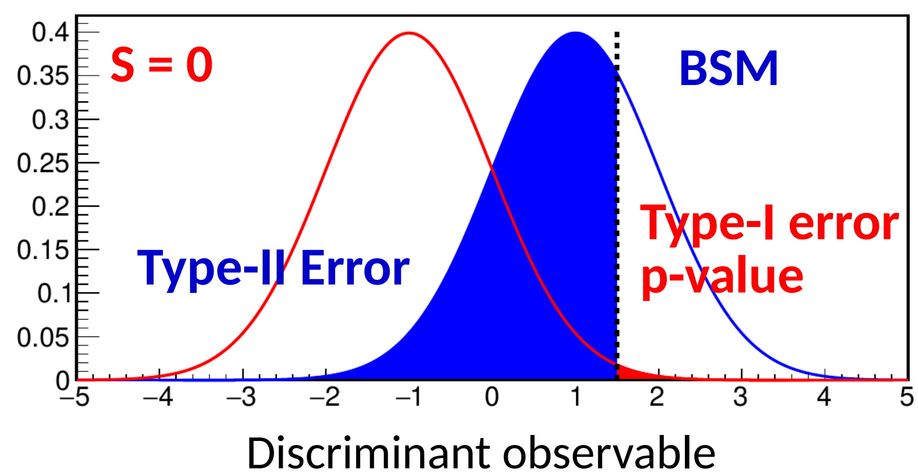
→ All curves monotonically decrease from (0,1) to (1,0)

→ Better discriminators more bent towards (1,1)

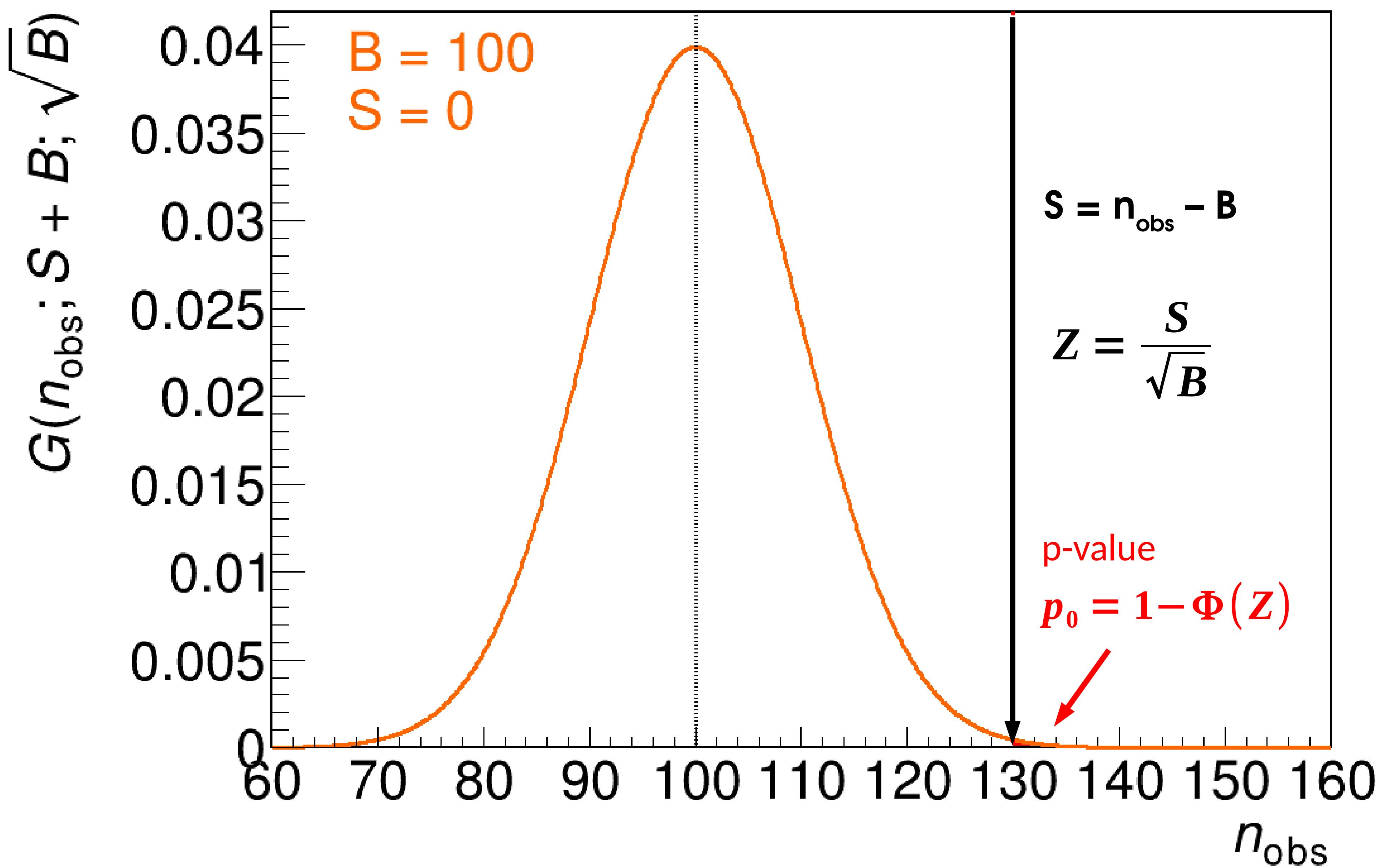


→ **Goal**: test that minimizes Type-II errors **for given level of Type-I error**.

→ Usually set predefined level of **acceptable Type-I error** (e.g. "5σ")



Discovery Testing in Gaussian counting



Hypothesis Testing with Likelihoods

Neyman-Pearson Lemma

When comparing two hypotheses H_0 and H_1, the optimal discriminator is the ***Likelihood ratio*** *(LR)*

$$\frac{L(H_0; data)}{L(H_1; data)}$$

e.g. $$\frac{L(S=0; data)}{L(S=5; data)}$$

Caveat: Strictly true only for *simple hypotheses* (no free parameters)

As for MLE, choose the hypothesis that is more likely **given the data we have**.

→ Always need an **alternate hypothesis** to test against the **null**.

→ **Minimizes Type-II uncertainties** for given level of Type-I uncertainties

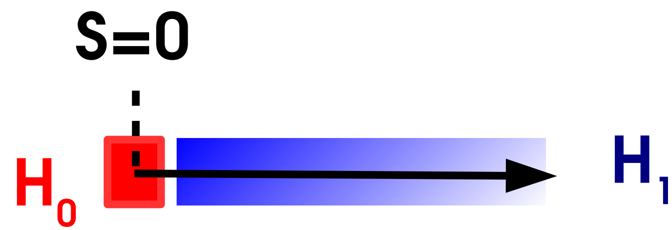
→ **In the following**: all tests based on LR, will focus on p-values (Type-I errors), trusting that Type-II errors are anyway as small as they can be...

Discovery: Test Statistic

Cowan, Cranmer, Gross & Vitells,
Eur.Phys.J.C71:1554,2011

Discovery :

- H_0 : **background only (S = 0)** against
- H_1: presence of a signal (**S > 0**)



→ For H_1, any S > 0 is possible, which to use ? **The one preferred by the data, Ŝ.**

⇒ Use Likelihood ratio: $\frac{L(S=0)}{L(\hat{S})}$

→ In fact use the **test statistic** $q_0 = -2\log\frac{L(S=0)}{L(\hat{S})}$

Note: for Ŝ < 0, set q_0=0 to reject negative signals ("*one-sided* test statistic")

Discovery p-value

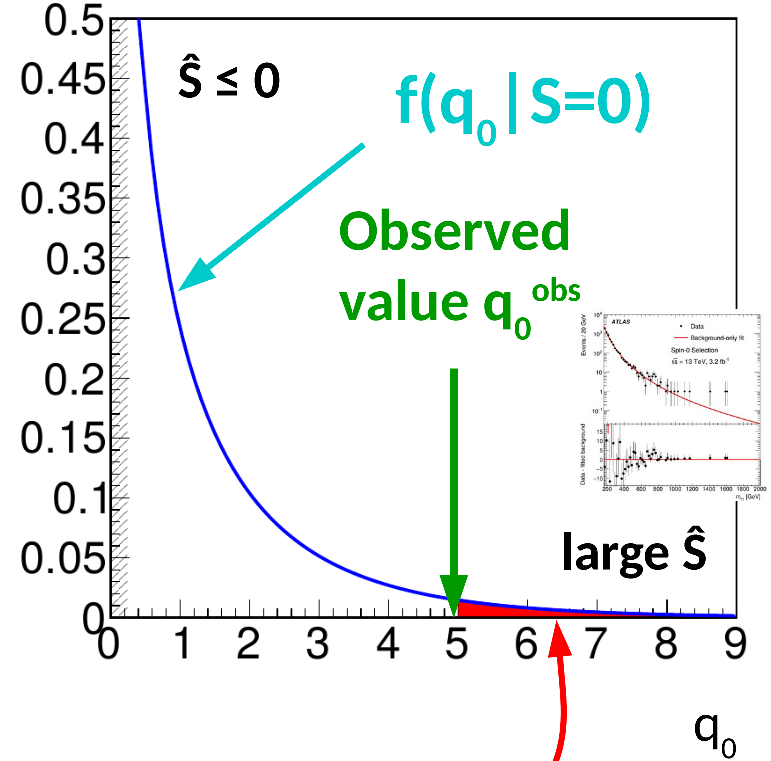
Large values of $-2\log\frac{L(S=0)}{L(\hat{S})}$ if:

⇒ observed $\hat{S}$ is far from 0

⇒ $H_0(S=0)$ ***disfavored*** compared to $H_1(S\neq0)$.

⇒ Large $\hat{S}$!

data prefer S = 0 ⟷ data prefer S > 0



Compute *p-value* in the tail of the distribution

to exclude H_0 (... and **claim a discovery!**)

$$p_0 = \int_{q_0^{obs}}^{\infty} f(q_0|S=0)\,dq_0$$

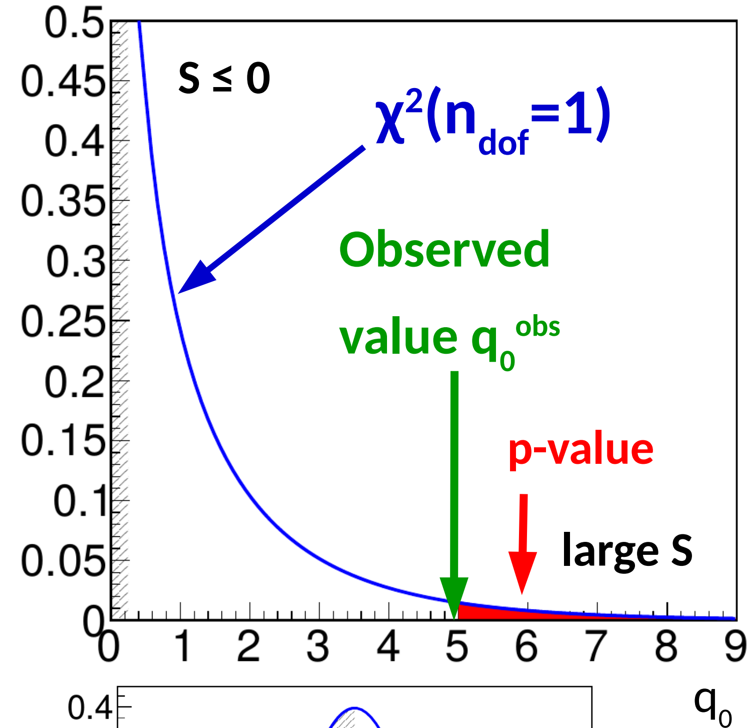
Need to know $f(q_0 \mid S=0)$, the distribution of the test statistic...

Asymptotic distribution of q_0

Cowan, Cranmer, Gross & Vitells
Eur.Phys.J.C71:1554,2011

Gaussian regime for $\hat{S}$ (e.g. large n_{evts}, Central-limit theorem) :

Wilks' Theorem: **q_0 distributed as $\chi^2(n_{par})$ for $S = 0$**



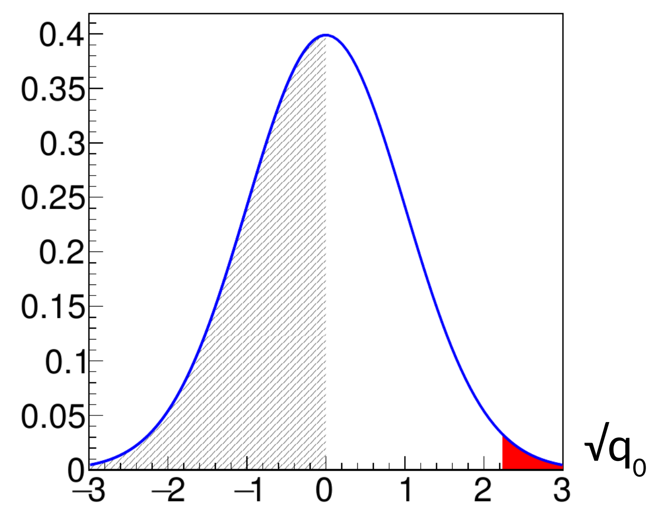
⇒ **$n_{par} = 1$: $\sqrt{q_0}$ is distributed as a Gaussian**

⇒ Can compute p-values from Gaussian quantiles

$$p_0 = 1 - \Phi\left(\sqrt{q_0}\right)$$

⇒ Even more simply, the significance is:

$$Z = \sqrt{q_0}$$



Typically works well already for for event counts of O(5) and above ⇒ Widely applicable

(*) 1-line "proof" : asymptotically L and S are Gaussian, so

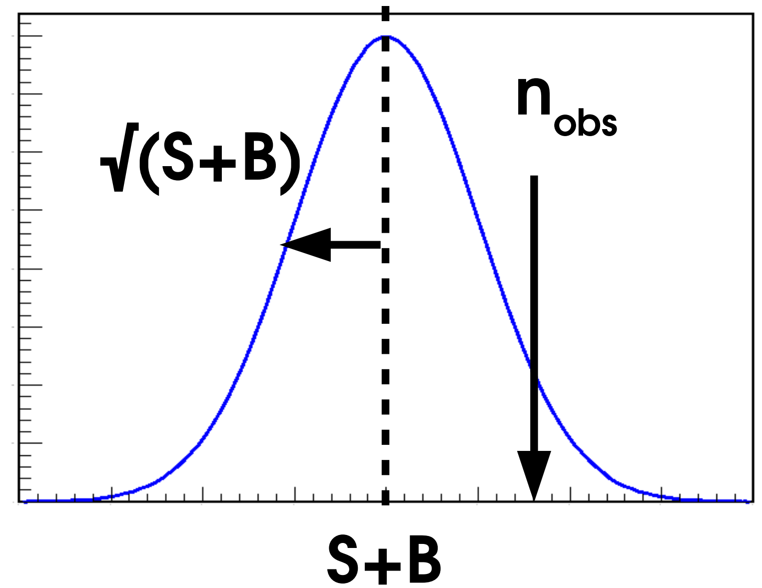
$$L(S) = \exp\left[-\frac{1}{2}\left(\frac{S-\hat{S}}{\sigma}\right)^2\right] \Rightarrow q_0 = \left(\frac{\hat{S}}{\sigma}\right)^2 \Rightarrow \sqrt{q_0} = \frac{\hat{S}}{\sigma} \sim G(0,1) \Rightarrow q_0 \sim \chi^2(n_{dof}=1)$$

Homework 1: Gaussian Counting

Count number of events n in data

→ Assume n large enough so process is Gaussian

→ Assume B is known, and we measure S



Likelihood :

$$L(S; n_{obs}) = e^{-\frac{1}{2}\left(\frac{n_{obs} - (S+B)}{\sqrt{S+B}}\right)^2}$$

→ Find the best-fit value (MLE) Ŝ for the signal

(can use λ = -2 log L instead of L for simplicity)

→ Find the expression of q_0 for Ŝ > 0.

→ Find the expression for the significance

$$Z = \frac{\hat{S}}{\sqrt{B}}$$

Homework 2: Poisson Counting

Same problem as Homework 1, but now ***not*** assuming Gaussian behavior:

$$L(S;n) = e^{-(S+B)}(S+B)^n$$

(Can remove the n! constant since we're only dealing with L ratios)

→ As before, compute $\hat{S}$, and q_0

→ Compute $Z = \sqrt{q_0}$, assuming asymptotic behavior

Solution:

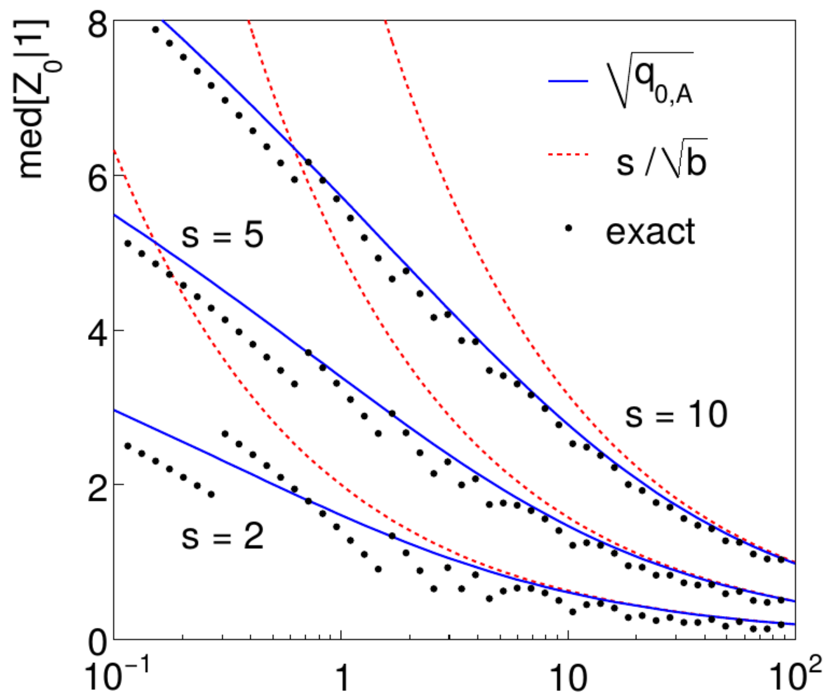
$$Z = \sqrt{2\left[(\hat{S}+B)\log\left(1+\frac{\hat{S}}{B}\right) - \hat{S}\right]}$$

Exact result can be obtained using

pseudo-experiments → close to $\sqrt{q_0}$ result

Asymptotic formulas justified by Gaussian regime, but remain valid even for small values of S+B (down to ~5 events!)

Eur.Phys.J.C71:1554,2011

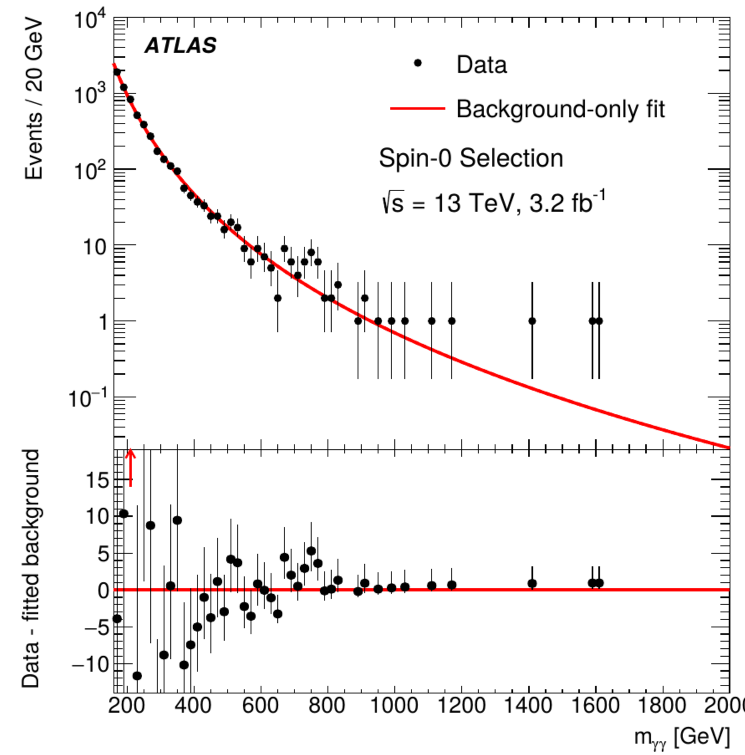


See G. Cowan's slides for the case with B uncertainty

Discovery Thresholds

Evidence : **3σ ⇔ p_0 = 0.3% ⇔ 1 chance in 300**

Discovery: **5σ ⇔ p_0 = 3 10^{-7} ⇔ 1 chance in 3.5M**



Why so high thresholds ? (from Louis Lyons):

- **Look-elsewhere effect**: searches typically cover multiple independent regions ⇒ Higher chance to have a fluctuation "somewhere"

  N_{trials} ~ 1000 : local 5σ ⇔ O(10^{-4}) more reasonable

- **Mismodeled systematics**: factor 2 error in syst-dominated analysis ⇒ factor 2 error on Z...

- **History**: 3σ and 4σ excesses do occur regularly, for the reasons above

Extraordinary claims require extraordinary evidence!

Highlights : Hypothesis Tests and Discovery

Given a PDF **P(data; μ)**, define likelihood **L(μ) = P(data; μ)**

To estimate a parameter, use the value $\hat{\mu}$ that maximizes L(μ) → best-fit value

To decide between hypotheses H_0 and H_1, use the **likelihood ratio** $\frac{L(H_0)}{L(H_1)}$

To test for **discovery**, use $q_0 = -2\log\frac{L(S=0)}{L(\hat{S})} \quad \hat{S} \geq 0$

For large enough datasets (n >~ 5), $Z = \sqrt{q_0}$

For a single **Gaussian** measurement, $Z = \frac{\hat{S}}{\sqrt{B}}$

For a single **Poisson** measurement, $Z = \sqrt{2\left[(\hat{S}+B)\log\left(1+\frac{\hat{S}}{B}\right) - \hat{S}\right]}$

Extra Slides

Categories

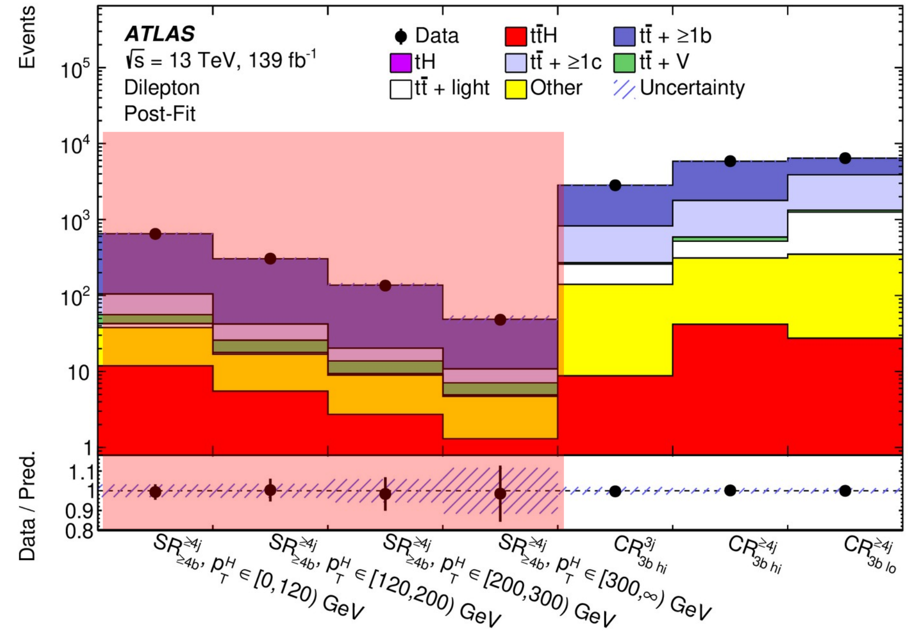
arXiv:2111.06712

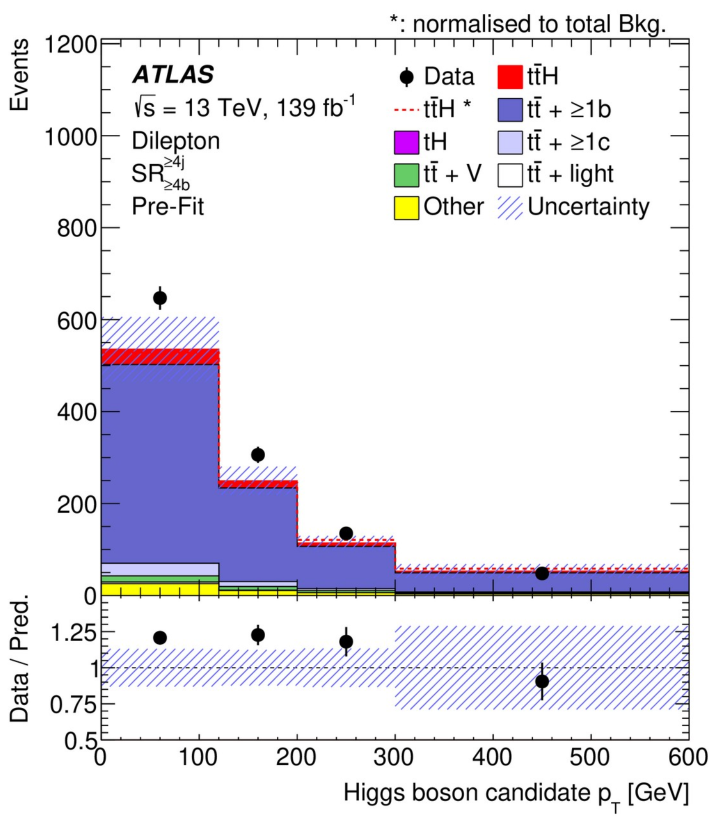
Multiple analysis regions often used.

→ Exploit better sensitivity in some regions

Here (ttH, H→bb analysis) **7** regions:

→ **4** *Signal Regions* (**SR**) split in p_T(Higgs)





Better sensitivity at high p_T

→ lower B backgrounds, higher S/B

Backgrounds levels from simulation here

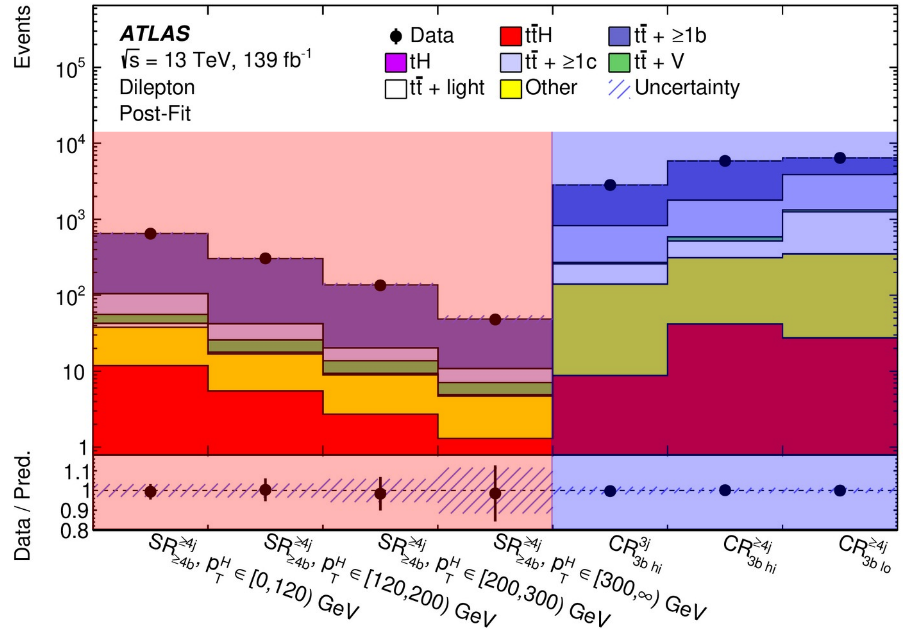
→ Large systematic uncertainties!

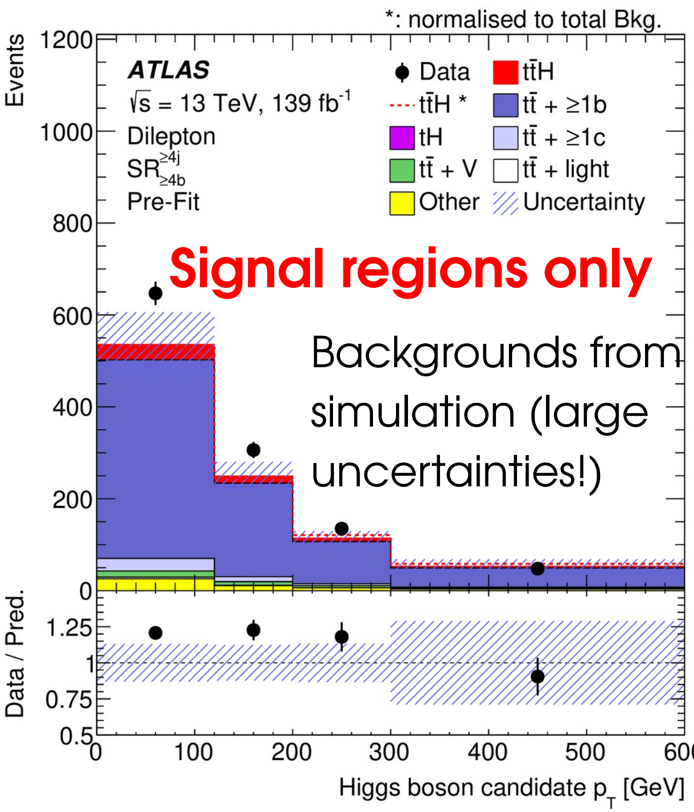
Categories

arXiv:2111.06712

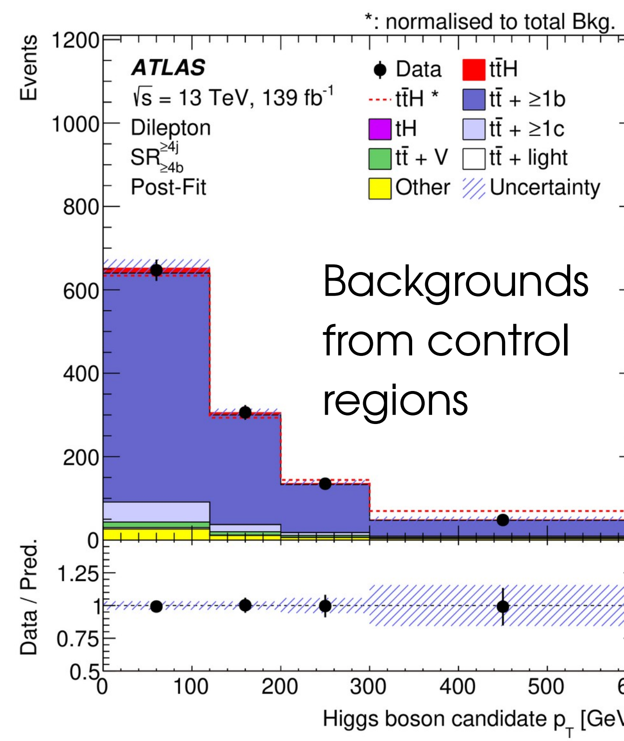
Multiple analysis regions often used.
→ Exploit better sensitivity in some regions
→ Constrain NPs: ***Control regions*** for bkgs

Here (ttH, H→bb analysis) **7** regions:
→ **4** *Signal Regions* (**SR**) split in p_T(Higgs)
→ **3** Background *Control Regions* (**CR**)





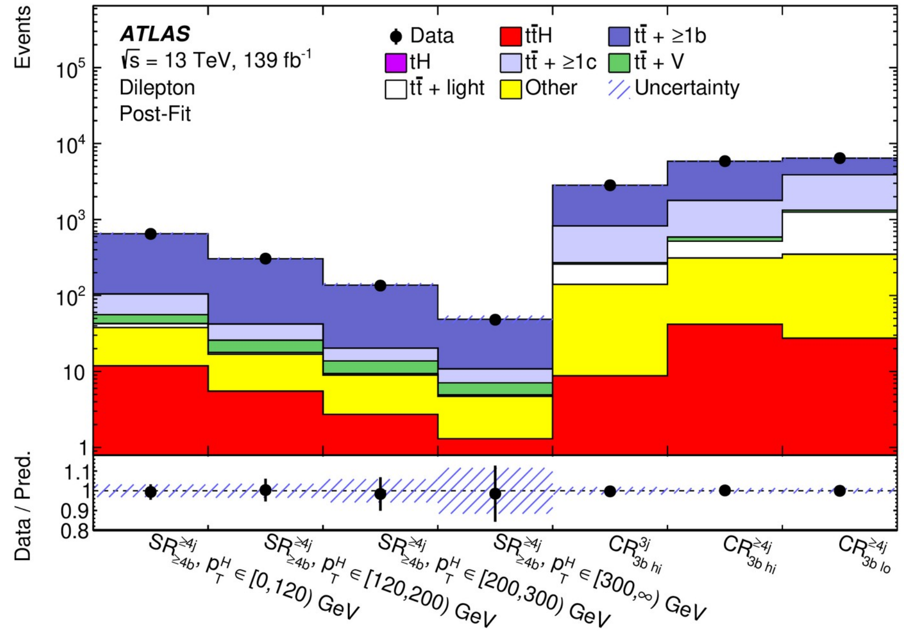
Include Background CRs



Signal + Bkg regions

Categories

arXiv:2111.06712



Multiple analysis regions often used.

→ Exploit better sensitivity in some regions

→ Constrain NPs: *Control regions* for bkgs

Here (ttH, H→bb analysis) 7 regions:

→ 4 *Signal Regions* (**SR**) split in p_T(Higgs)

→ 3 Background *Control Regions* (**CR**)

⇒ ***Combined PDF*** :

PDF for category k

$$P\left(S, B; \{n_i^{(k)}\}_{i=1\ldots n_{\text{evts}}^{(k)}}^{k=1\ldots n_{\text{cats}}}\right) = \prod_{k=1}^{n_{\text{cats}}} P_k\left(S, B; \{n_i^{(k)}\}_{i=1\ldots n_{\text{evts}}^{(k)}}\right)$$

No overlaps between categories ⇒ No statistical correlations

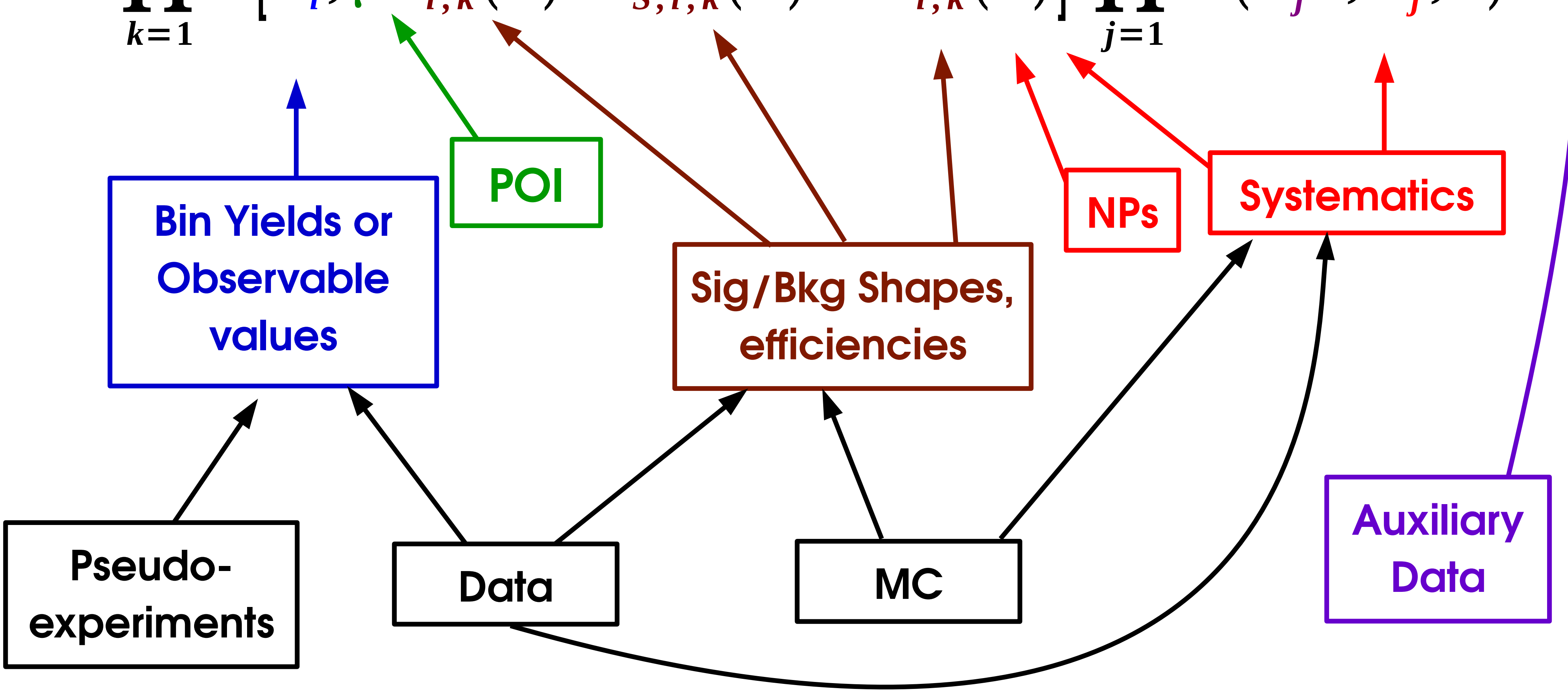
⇒ can simply take product of individual PDFs.

Multiple categories allows to **constrain nuisance parameters** (e.g. **B**)

Counting model, the full version

$$P\left(\mu,\{\theta_j\}_{j=1\ldots n_{NP}};\{n_i^{(k)}\}_{i=1\ldots n_{data}^{(k)}}^{k=1\ldots n_{cat}},\{\theta_j^{obs}\}_{j=1..n_{NP}}\right)=$$

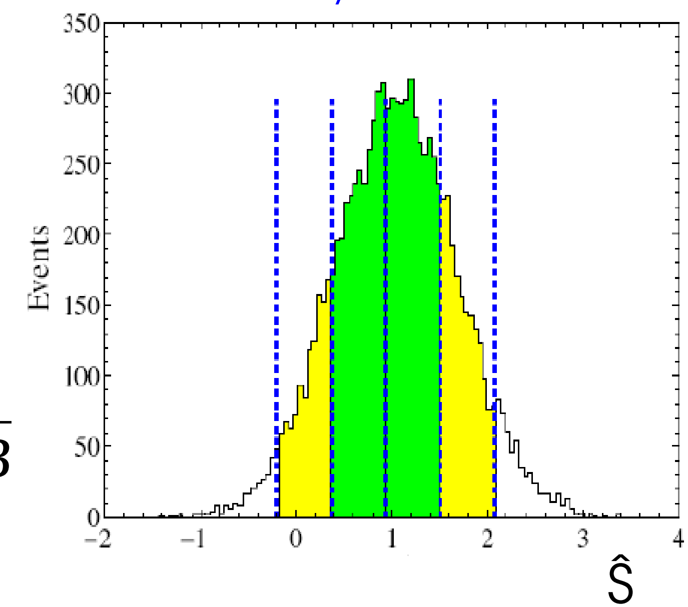
$$\prod_{k=1}^{n_{cats}} P\left[n_i;\mu\,\epsilon_{i,k}(\vec{\theta})\;N_{S,i,k}(\vec{\theta})+B_{i,k}(\vec{\theta})\right]\prod_{j=1}^{n_{syst}} G(\theta_j^{obs};\theta_j;1)$$



× number of categories!

CL_s : Gaussian Bands

Eur.Phys.J.C71:1554,2011



Usual Gaussian counting example with known B:

95% CL_s upper limit on S:

$$S_{up} = \hat{S} + \left[\Phi^{-1}\left(1 - 0.05\,\Phi\left(\hat{S}/\sigma_S\right)\right)\right]\sigma_S$$

with

$$\sigma_S = \sqrt{B}$$

Compute expected bands for S=0:

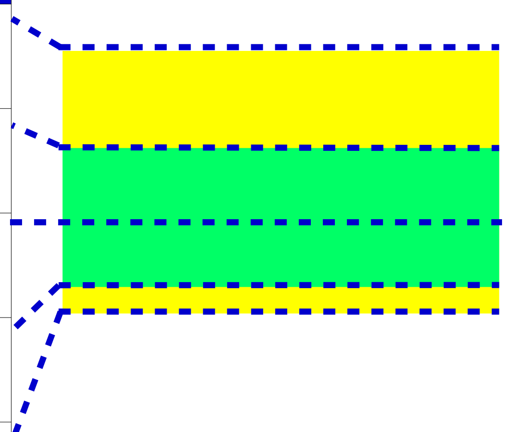
→ **Asimov dataset ⇔ $\hat{S} = 0$** :

$$S^0_{\text{up,exp}} = 1.96\,\sigma_S$$

→ **± nσ bands**:

$$S^{\pm n}_{\text{up,exp}} = \left(\pm n + \left[1 - \Phi^{-1}\left(0.05\,\Phi(\mp n)\right)\right]\right)\sigma_S$$

n	$S_{exp}^{\pm n}$ / √B
+2	3.66
+1	2.72
0	1.96
-1	1.41
-2	1.05



CLs :
- Positive bands somewhat reduced,
- Negative ones more so

Band width from $\sigma^2_{S,A} = \dfrac{S^2}{q_S(\text{Asimov})}$ depends on S, for non-Gaussian cases,different values for each band...

Comparison with LEP/TeVatron definitions

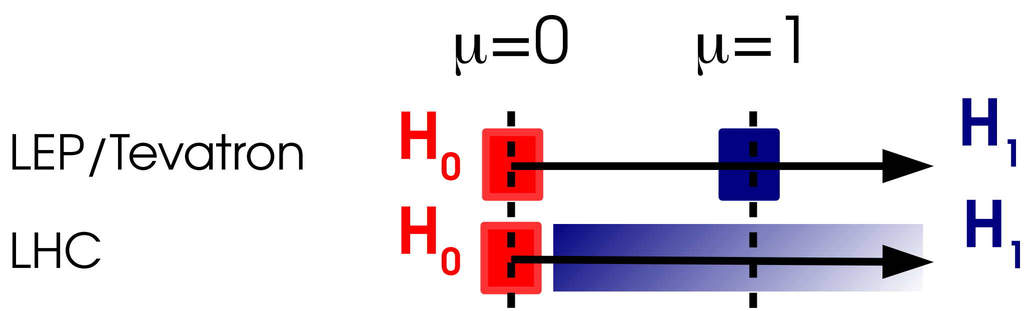
Likelihood ratios are not a new idea:

- **LEP**: Simple LR with NPs from MC
  - Compare μ=0 and μ=1
- **Tevatron**: PLR with profiled NPs

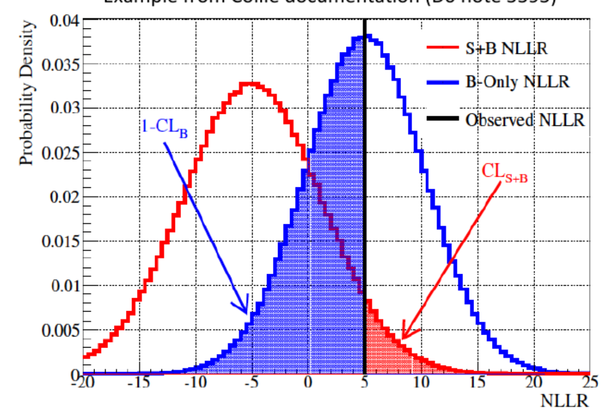
$$q_{LEP} = -2\log\frac{L(\mu=0,\tilde{\theta})}{L(\mu=1,\tilde{\theta})}$$

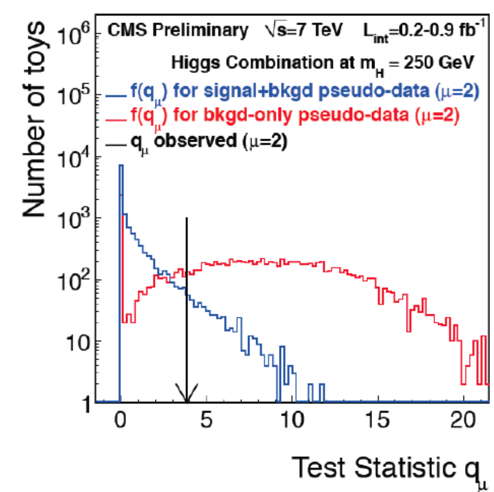
$$q_{Tevatron} = -2\log\frac{L(\mu=0,\hat{\hat{\theta}}_0)}{L(\mu=1,\hat{\hat{\theta}}_1)}$$

Both compare to **μ=1** instead of best-fit **$\hat{\mu}$**



Example from Collie documentation (D0 note 5595)





→ Asymptotically:

- **LEP/Tevaton**: q linear in μ ⇒ **~Gaussian**
- **LHC**: q quadratic in μ ⇒ **~χ2**

→ Still use TeVatron-style for discrete cases

Andrey Korytov, EPS 2011

Wilks' Theorem

Cowan, Cranmer, Gross & Vitells
Eur.Phys.J.C71:1554,2011

To test the $S=S_0$ hypothesis, consider

$$t(S_0) = -2\log\frac{L(S=S_0)}{L(\hat{S})}$$

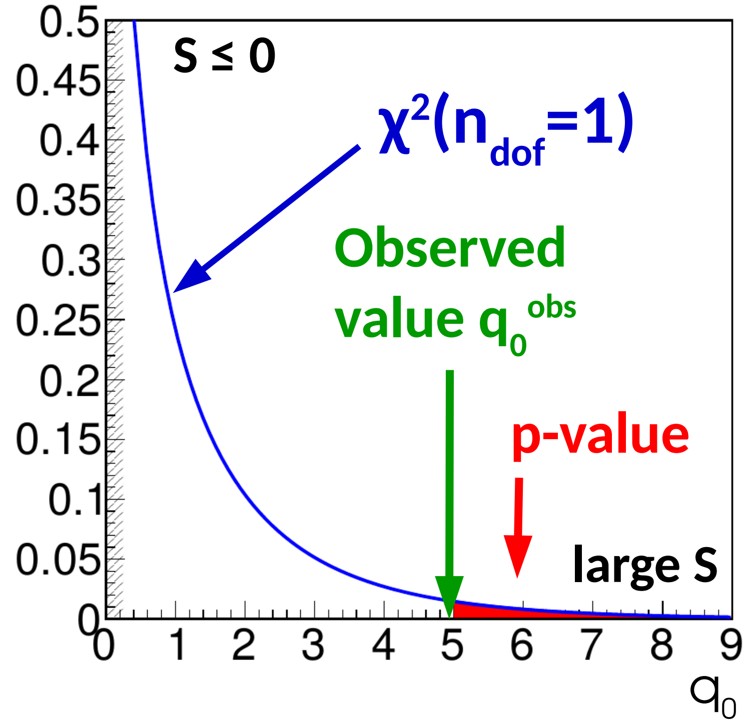
→ Assume **Gaussian regime** (e.g. large n_{evts}, Central-limit theorem) : then:

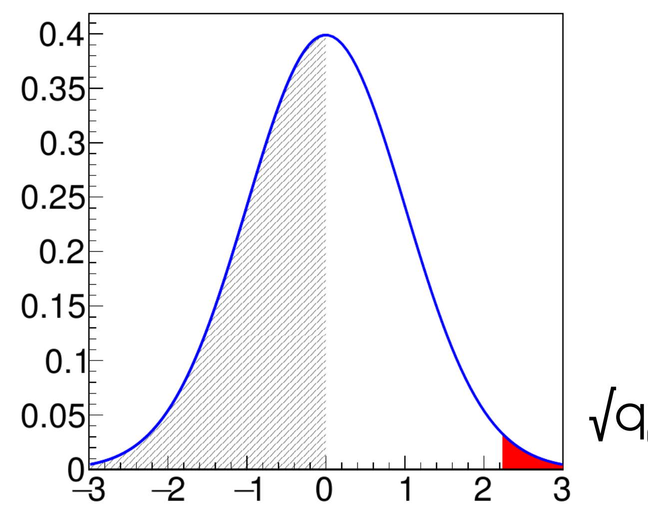
Wilk's Theorem: **$t(S_0)$ is distributed as a χ^2**

under $S=S_0$: $f\left(t_{S_0} \mid S=S_0\right) = f_{\chi^2(n_{dof}=1)}\left(t_{S_0}\right)$

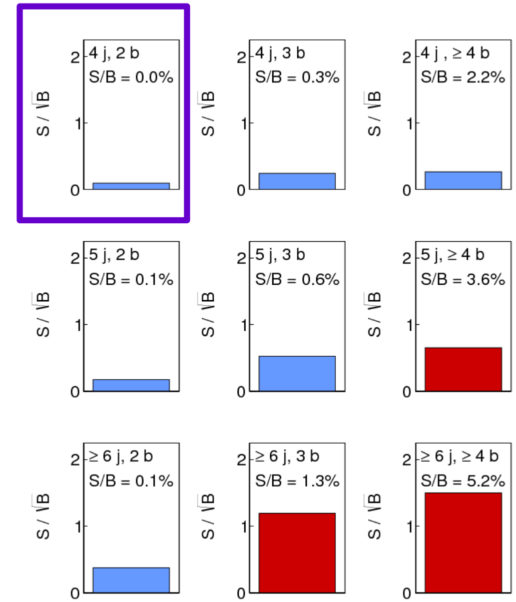
⇒ In particular, the significance is:

$$Z = \sqrt{q_0}$$





Profiling Example: ttH→bb

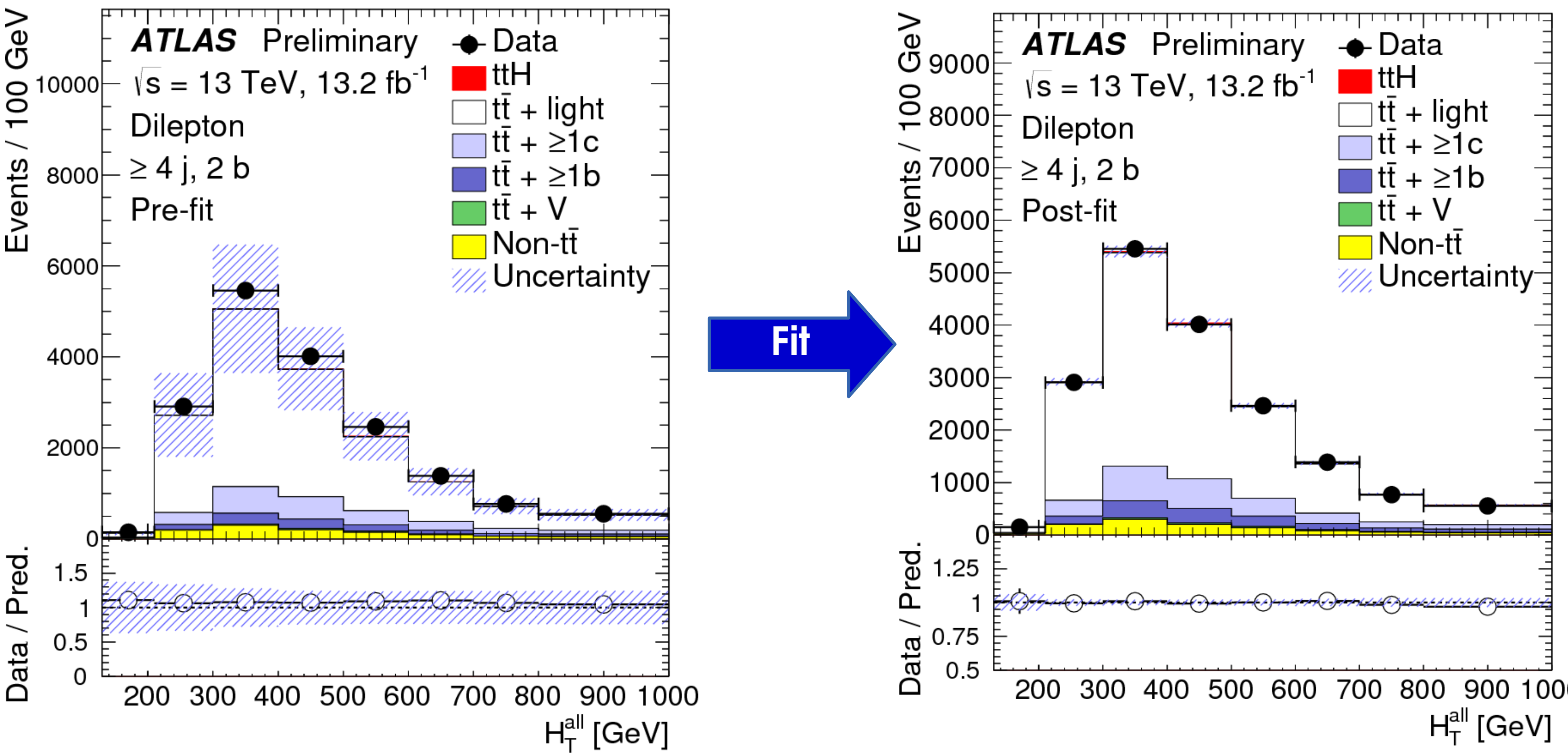


Analysis uses low-S/B categories to constrain backgrounds.

→ **Reduction in large uncertainties on tt bkg**

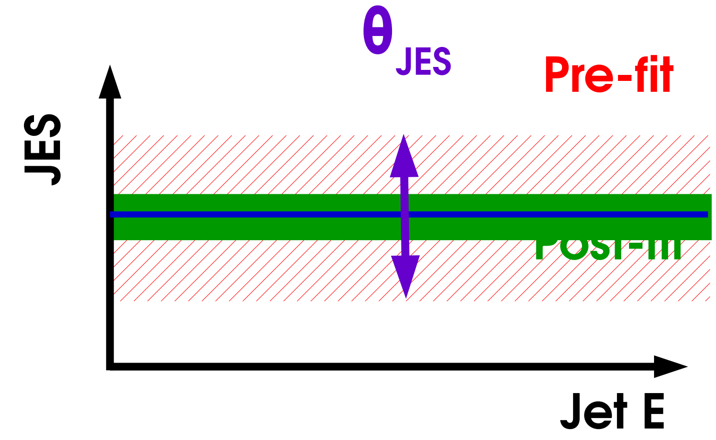
→ **Propagates to the high-S/B categories** through the statistical modeling

⇒ **Care needed in the propagation** (e.g. different kinematic regimes)



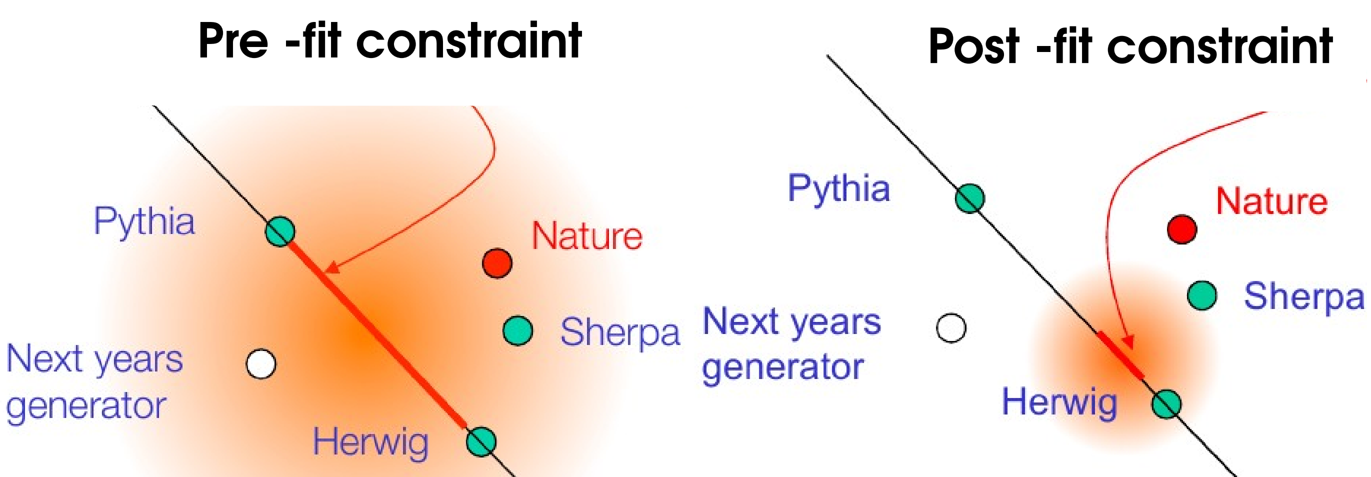
ATLAS-CONF-2016-08

Profiling Issues



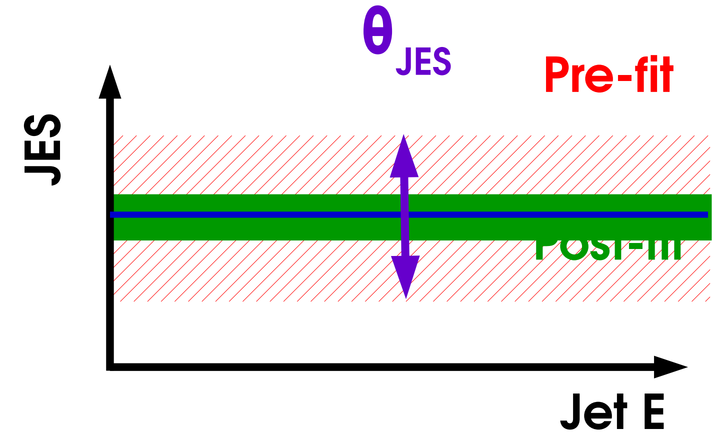
Too simple modeling can have unintended effects
→ e.g. single Jet E scale parameter:
⇒ Low-E jets calibrate high-E jets – intended ?

***Two-point* uncertainties**:
→ Interpolation may not cover full configuration space, can lead to too-strong constraints



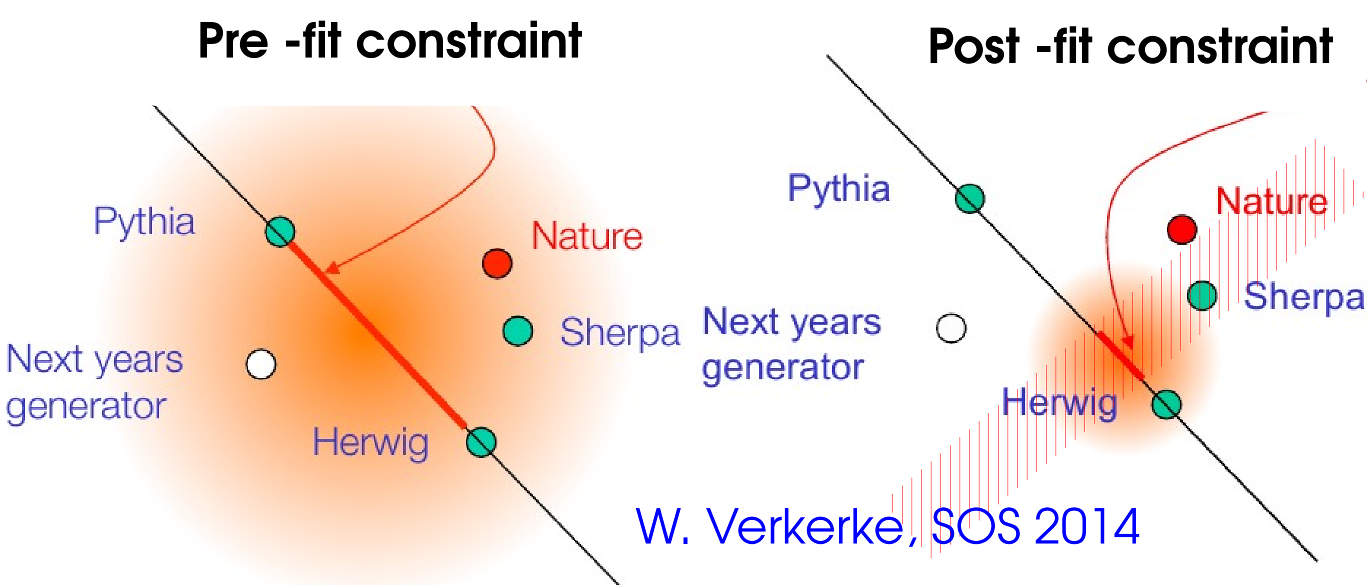
W. Verkerke, SOS 2014

Profiling Issues



Too simple modeling can have unintended effects
→ e.g. single Jet E scale parameter:
⇒ Low-E jets calibrate high-E jets – intended ?

***Two-point* uncertainties**:
→ Interpolation may not cover full configuration space, can lead to too-strong constraints



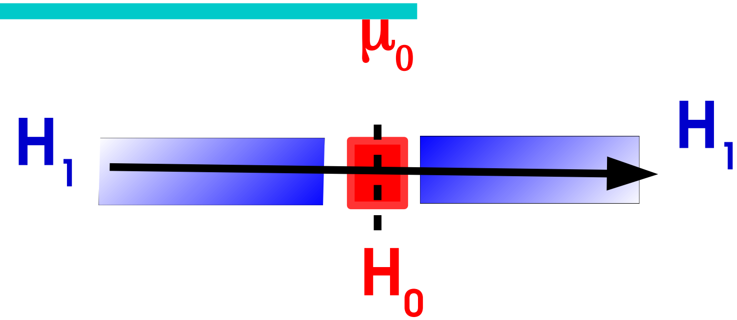
W. Verkerke, SOS 2014

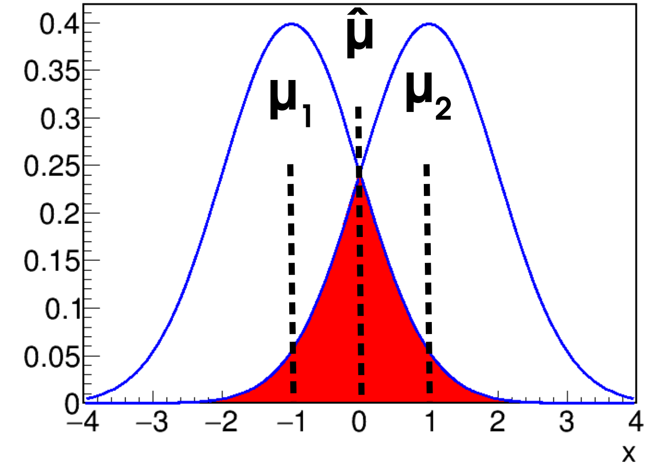
Test Statistics for Limit-Setting

Interval :

$H_0 : \mu = \mu_0$

$H_1 : \mu \neq \mu_0$





Try to exclude μ values away from $\hat{\mu}$.

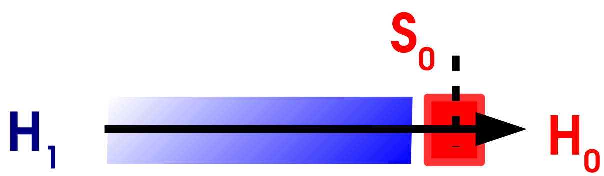
$$t(\mu_0) = -2\log\frac{L(\mu=\mu_0)}{L(\hat{\mu})}$$

"**Two-sided**" test

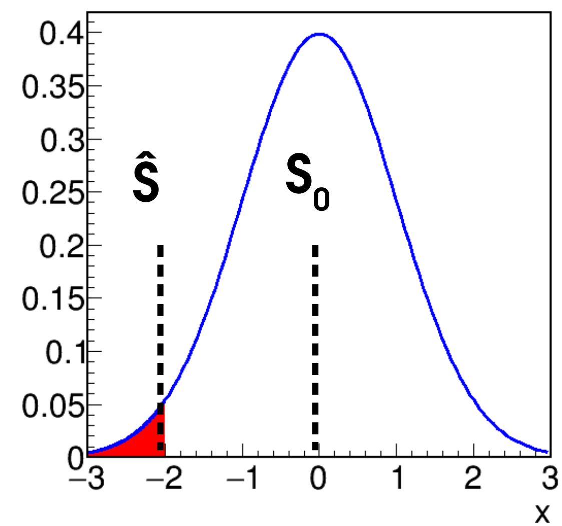
Limit-setting

$H_0 : S = S_0$

$H_1 : S < S_0$



$$q(S_0) = \begin{cases} -2\log\frac{L(S=S_0)}{L(\hat{S})} & S_0 > \hat{S} \\ 0 & S_0 \leq \hat{S} \end{cases}$$



Try to exclude values of S that are above $\hat{S}$.

⇒ "**One-sided**" test : only interested in excluding above

Discovery is also one-sided, for S>0 !

Hands-on session

The hands-on session will be based on **jupyter notebooks** built using the **`numpy/scipy/pyplot`** stack.

If you have a computer with you, **please install** anaconda as this provides a consistent installation of python, JupyterLab, etc.

→ Alternatively, you can also install JupyterLab as a standalone package.

→ Another solution is to run the notebooks on the public jupyter servers at mybinder.org. This will probably be slower but avoids a local install.

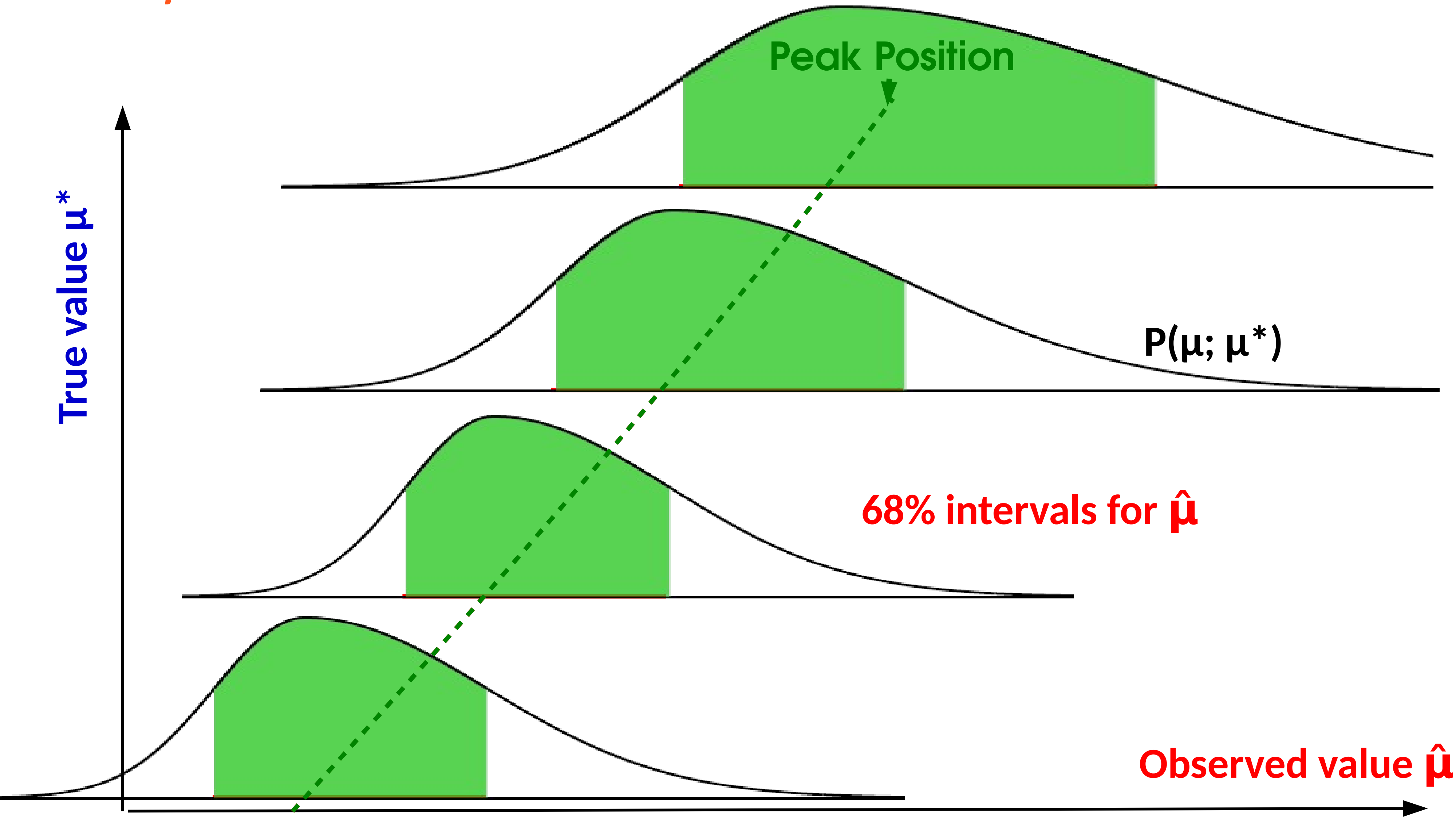
Lecture 1		notebook [solutions]	binder [solutions]
Lecture 1	Lecture Notes	notebook	binder
Lecture 2	Lecture notes	notebook	binder

The **warmup** item includes material that will not be covered in detail in the class, as well as an introduction to the notebooks. **Please have a look before the beginning of the classes** if you are unfamiliar with any of this.

Neyman Construction

General case: build 1σ intervals of observed values for each true value

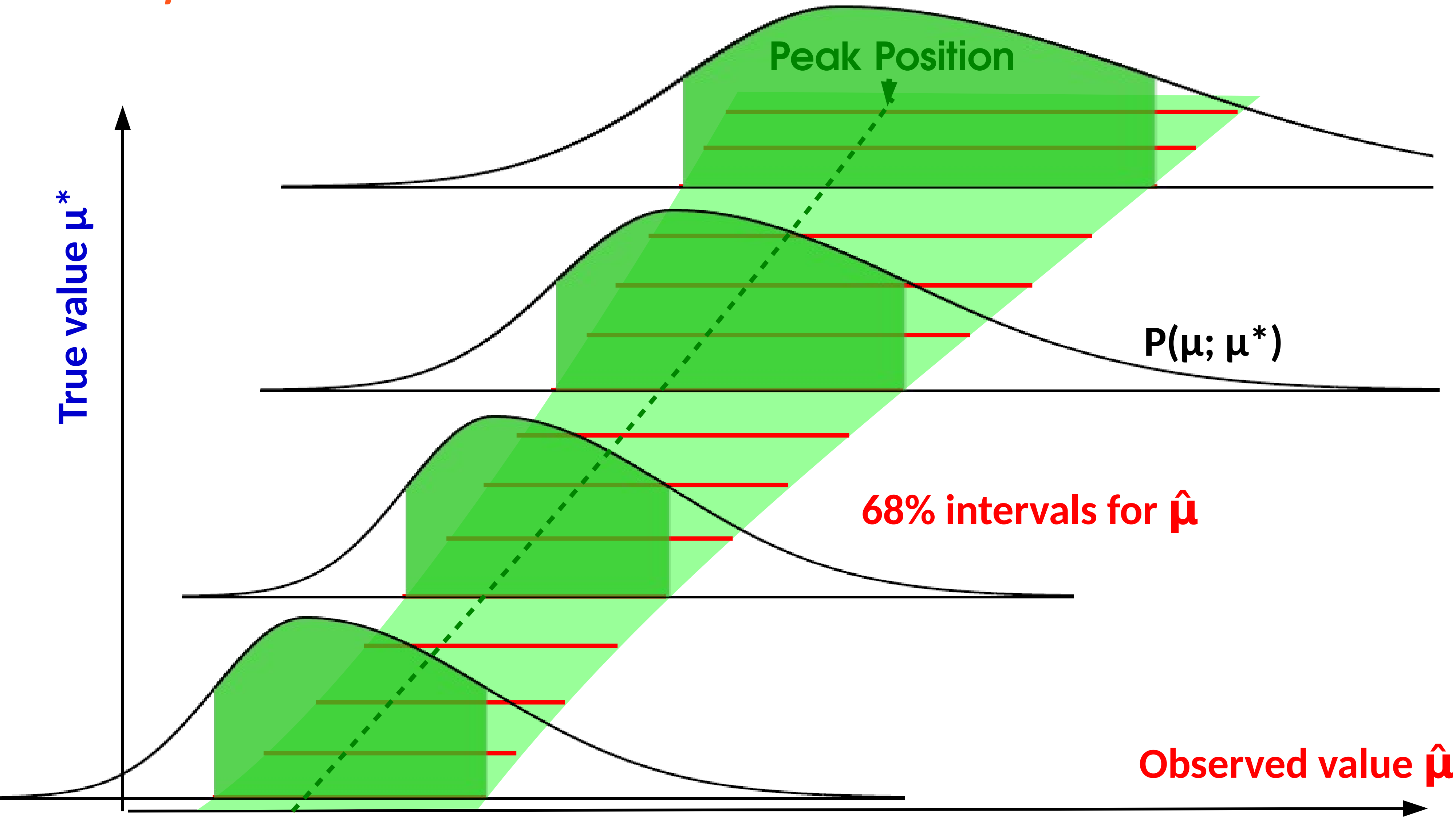
⇒ Confidence belt



Neyman Construction

General case: build 1σ intervals of observed values for each true value

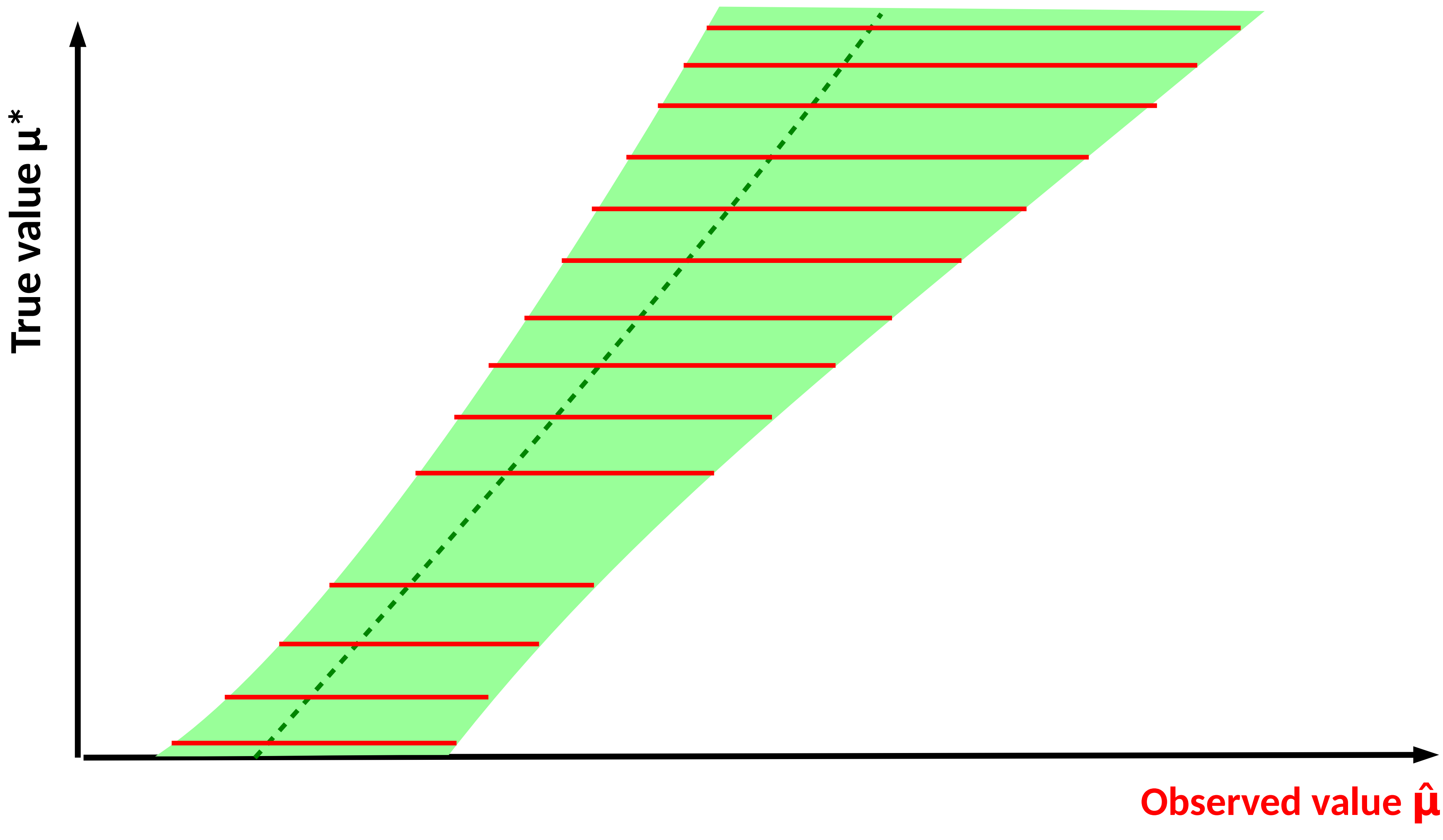
⇒ Confidence belt



Inversion using the Confidence Belt

General case: Intersect belt with given $\hat{\mu}$, get $P(\hat{\mu} - \sigma_\mu^- < \mu^* < \hat{\mu} + \sigma_\mu^+) = 68\,\%$

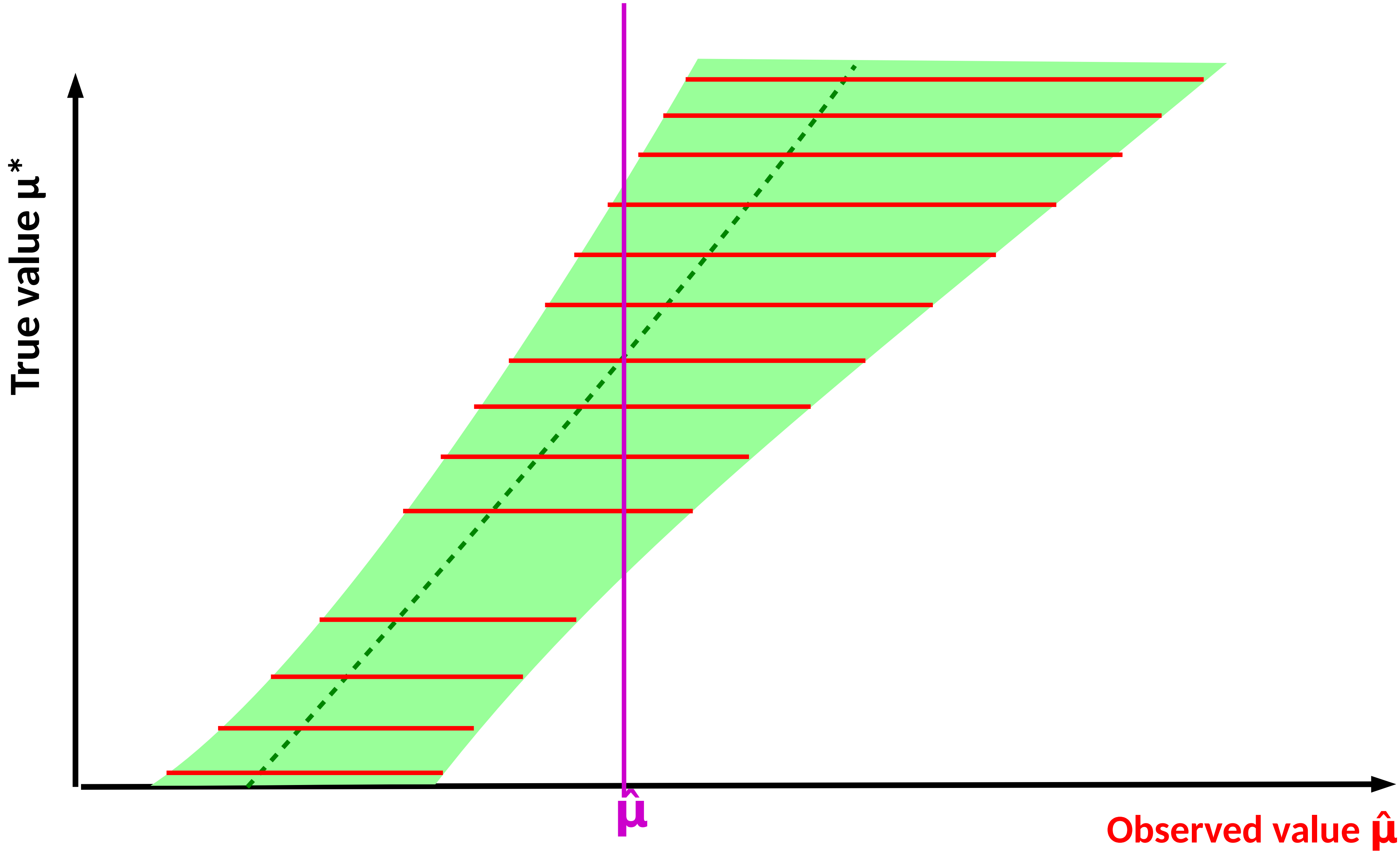
→ Same as before for Gaussian, works also when $P(\mu^{obs}|\mu)$ varies with μ.



Inversion using the Confidence Belt

General case: Intersect belt with given $\hat{\mu}$, get $P(\hat{\mu} - \sigma_\mu^- < \mu^* < \hat{\mu} + \sigma_\mu^+) = 68\,\%$

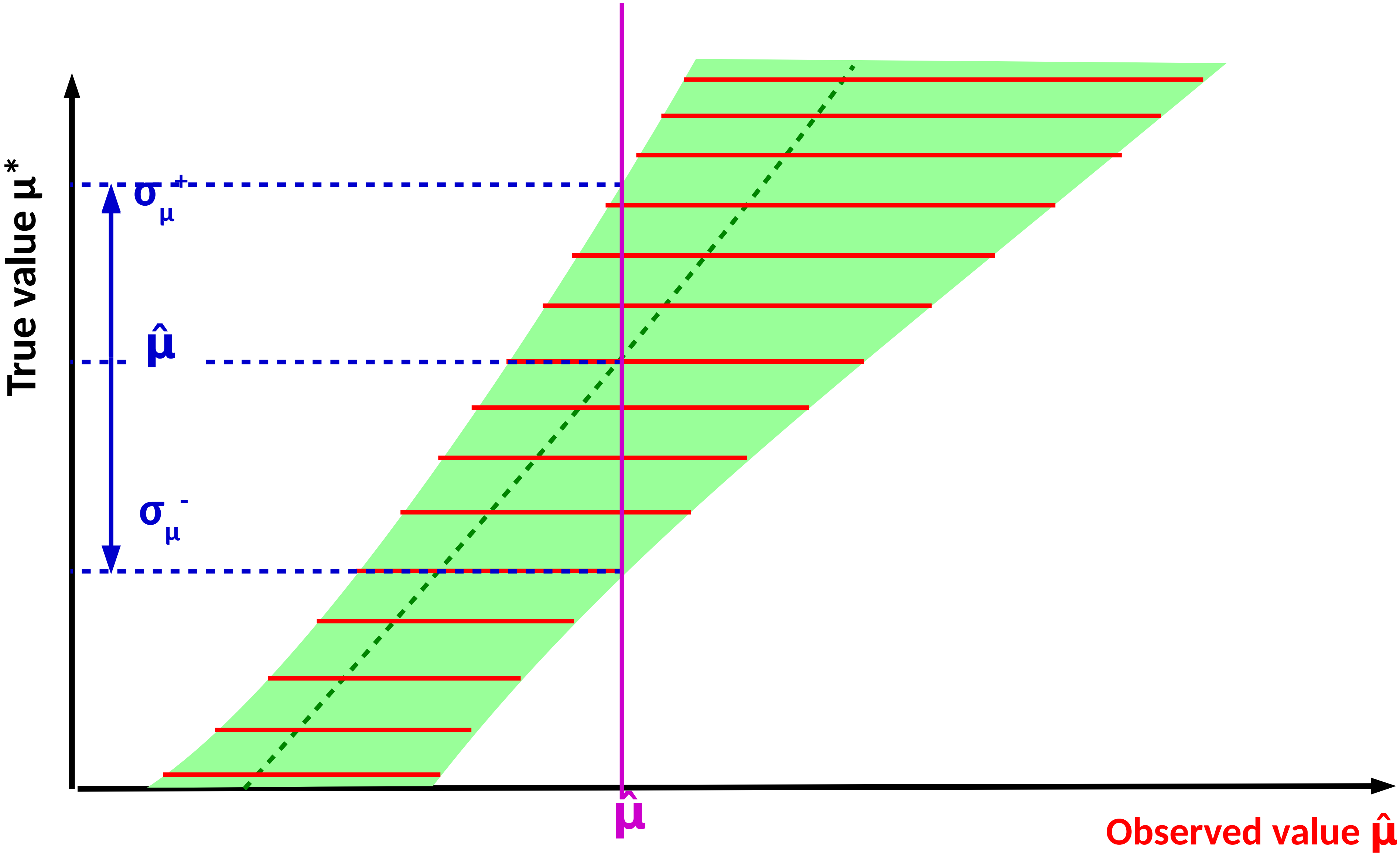
→ Same as before for Gaussian, works also when $P(\mu^{obs}|\mu)$ varies with μ.



Inversion using the Confidence Belt

General case: Intersect belt with given $\hat{\mu}$, get $P\left(\hat{\mu} - \sigma_\mu^- < \mu^* < \hat{\mu} + \sigma_\mu^+\right) = 68\,\%$

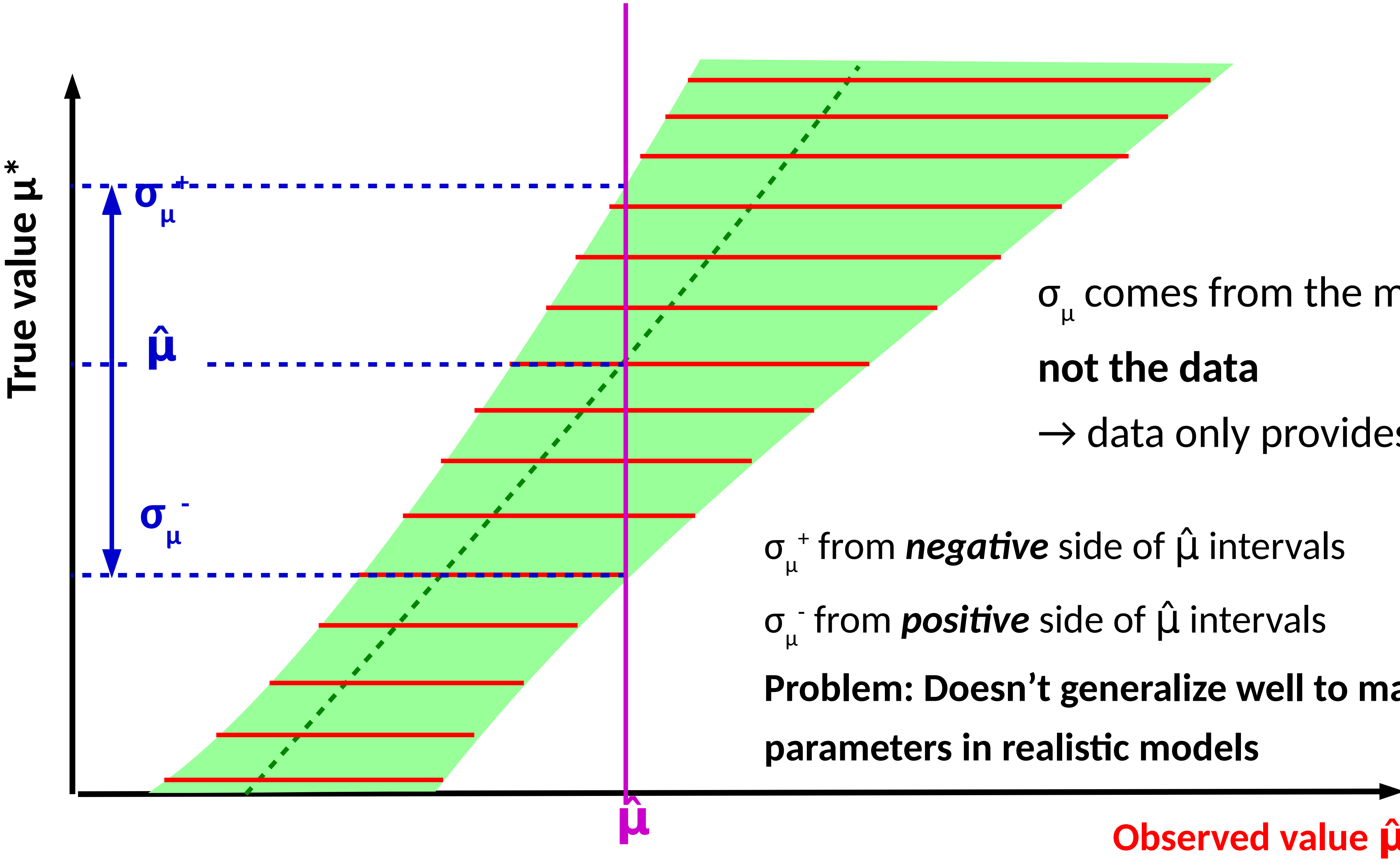
→ Same as before for Gaussian, works also when $P(\mu^{obs}|\mu)$ varies with μ.



Inversion using the Confidence Belt

General case: Intersect belt with given $\hat{\mu}$, get $P\left(\hat{\mu} - \sigma_\mu^- < \mu^* < \hat{\mu} + \sigma_\mu^+\right) = 68\,\%$

→ Same as before for Gaussian, works also when $P(\mu^{obs}|\mu)$ varies with μ.



σ_μ comes from the model,

not the data

→ data only provides $\hat{\mu}$.

σ_μ^+ from ***negative*** side of $\hat{\mu}$ intervals

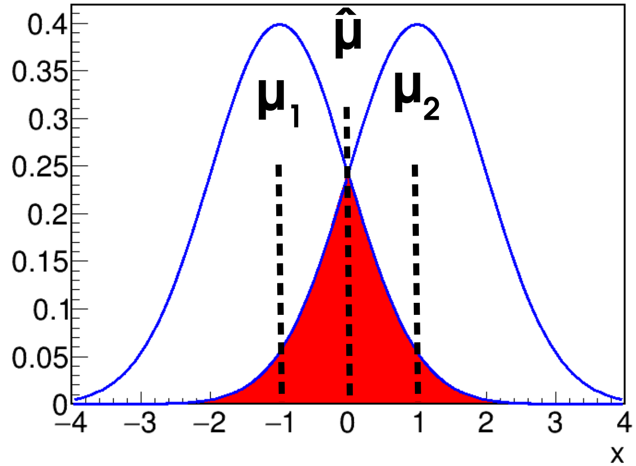
σ_μ^- from ***positive*** side of $\hat{\mu}$ intervals

Problem: Doesn't generalize well to many parameters in realistic models

Test Statistics for Limit-Setting

Confidence Interval :

Try to exclude μ values away from $\hat{\mu}$.

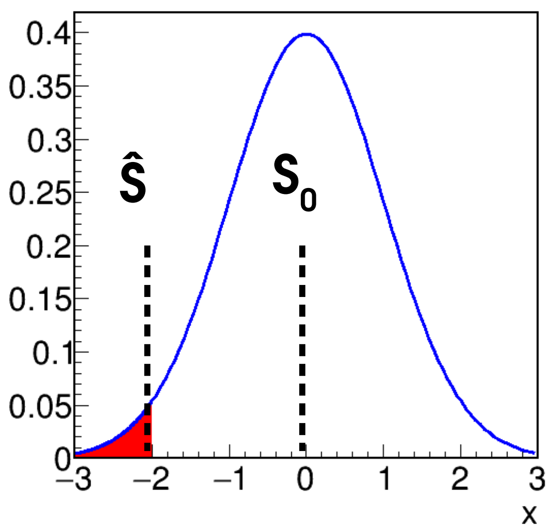


$$t(\mu_0) = -2\log\frac{L(\mu=\mu_0)}{L(\hat{\mu})}$$

"**Two-sided**" test

Limit-setting

Try to exclude values of S that are above $\hat{S}$.



$$q(S_0) = \begin{cases} -2\log\dfrac{L(S=S_0)}{L(\hat{S})} & S_0 > \hat{S} \\ 0 & S_0 \leq \hat{S} \end{cases}$$

"One-sided" test : only interested in excluding above

Discovery was also one-sided, for S>0

Inversion : Getting the limit for a given CL

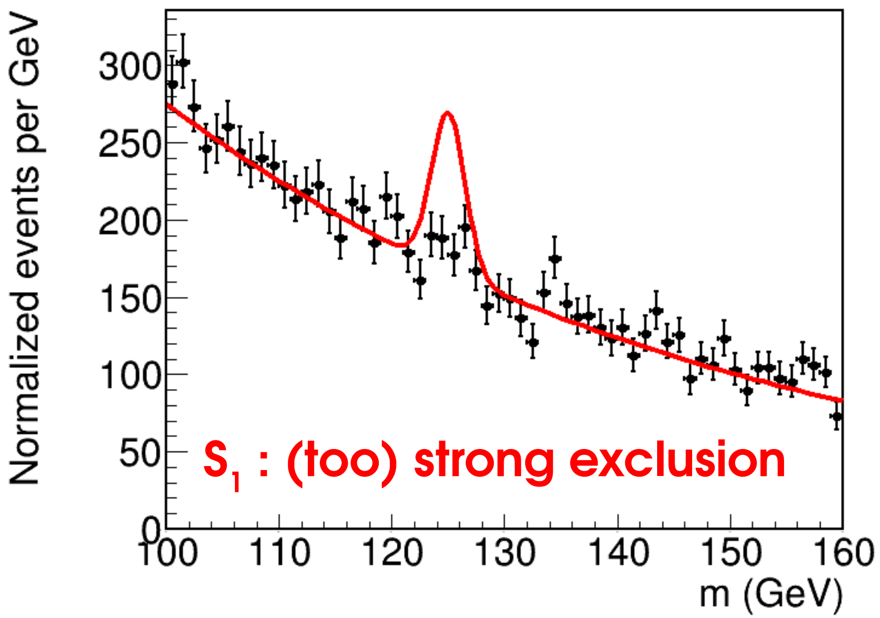
Procedure:

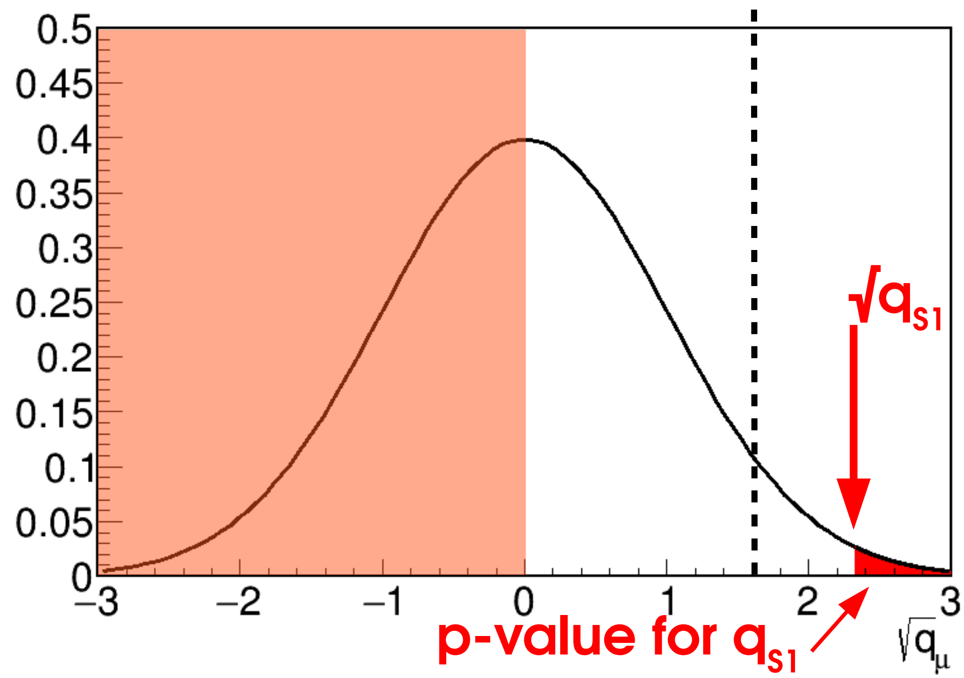
→ Compute $q(S_0)$ for some S_0,
get the exclusion p-value $p(S_0)$.

Asymptotics: $$p(S_0) = 1 - \Phi\left(\sqrt{q(S_0)}\right)$$

CL	p	Region
90%	10%	$\sqrt{q(S)} > 1.28$
95%	5%	$\sqrt{q(S)} > 1.64$
99%	1%	$\sqrt{q(S)} > 2.33$

→ **Adjust S_0** to get the desired exclusion
Asymptotics: need $\sqrt{q(S_{95})} = 1.64$ for **95% CL**





Inversion : Getting the limit for a given CL

Procedure:

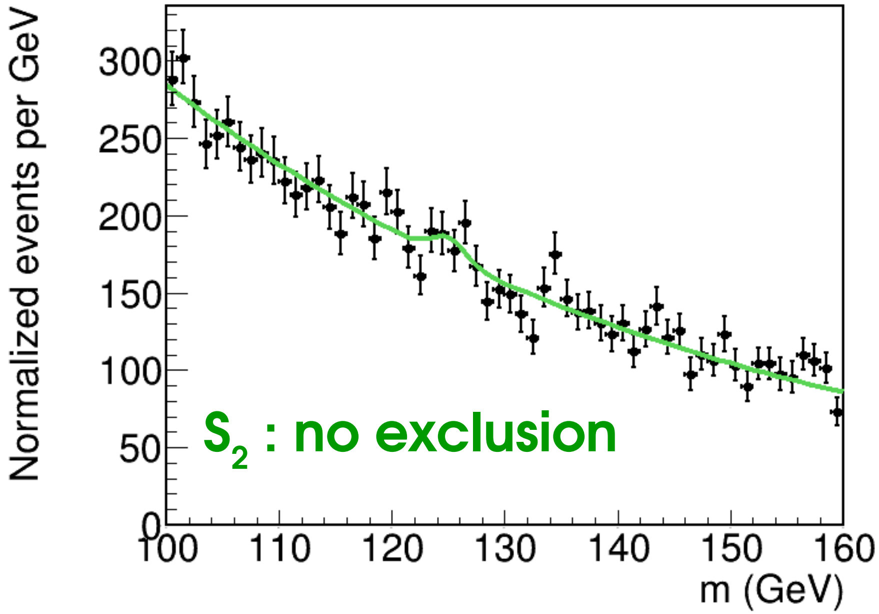
→ Compute $q(S_0)$ for some S_0, get the exclusion p-value $p(S_0)$.

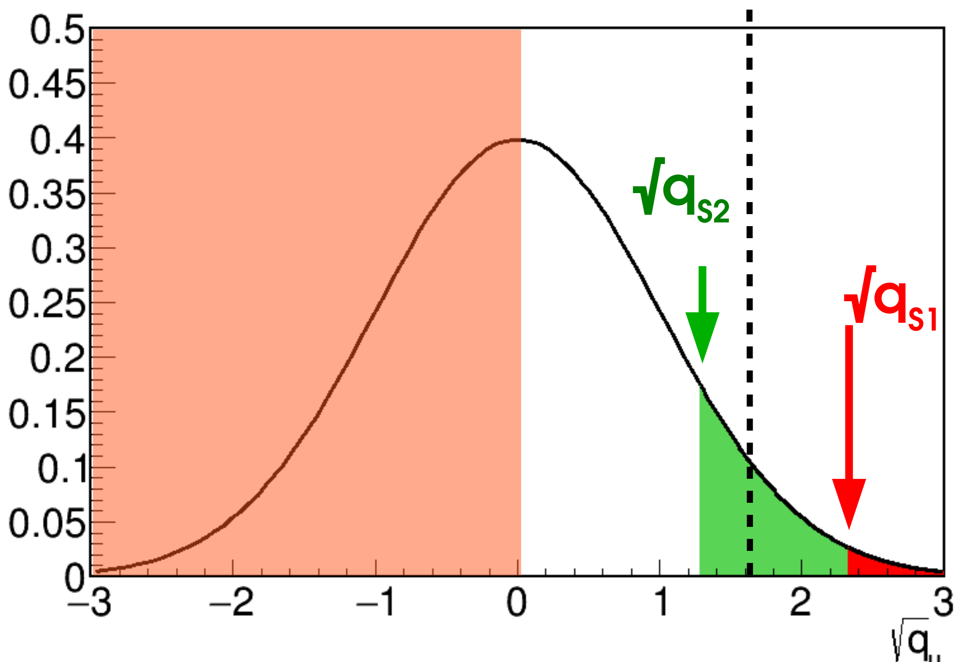
Asymptotics: $$p(S_0) = 1 - \Phi\left(\sqrt{q(S_0)}\right)$$

CL	p	Region
90%	10%	$\sqrt{q(S)} > 1.28$
95%	5%	$\sqrt{q(S)} > 1.64$
99%	1%	$\sqrt{q(S)} > 2.33$

→ **Adjust S_0** to get the desired exclusion

Asymptotics: need $\sqrt{q(S_{95})} = 1.64$ for **95% CL**





Inversion : Getting the limit for a given CL

Procedure:

→ Compute $q(S_0)$ for some S_0, get the exclusion p-value $p(S_0)$.

Asymptotics: $$p(S_0) = 1 - \Phi\left(\sqrt{q(S_0)}\right)$$

CL	p	Region
90%	10%	$\sqrt{q(S)} > 1.28$
95%	5%	$\sqrt{q(S)} > 1.64$
99%	1%	$\sqrt{q(S)} > 2.33$

→ **Adjust S_0** to get the desired exclusion

Asymptotics: need $\sqrt{q(S_{95})} = 1.64$ for **95% CL**