From detection to high-level reconstruction

A statistical view on gamma-ray sources

2024.10.15, Č Astronomy Data School

Jonathan Biteau

I. Gamma-ray astronomy as a counting experiment

Concepts: *DL3-DL4-DL5, forward-folding, binomial / Poisson / Gauss, PMF / PDF / CDF / SF*

II. Statistical inference

Concepts: *Bayesian & frequentist, marginalizing & profiling, parameter estimation, (saturated) deviance*

III. Model selection

Concepts: *goodness of fit, residuals & pulls, under-/over-fitting,*
local & global p-values (trial factors / look-elsewhere), likelihood ratio test, significance level, Li & Ma

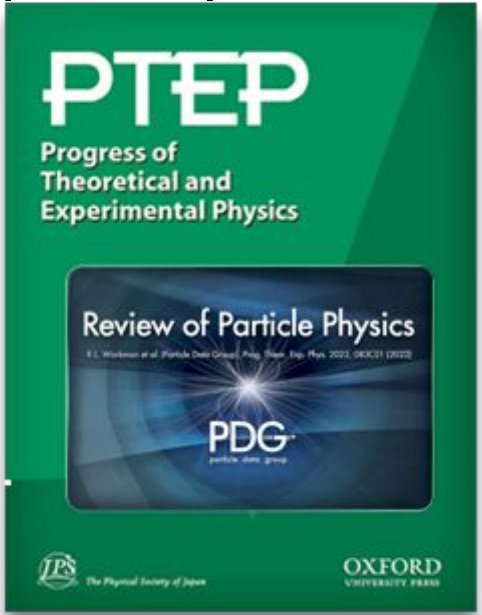
IV. Dealing with uncertainties

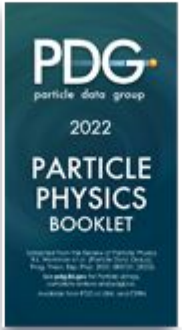
Concepts: *covariance, decorrelation energy, bowtie, systematic uncertainties, bracketing / marginalizing*

V. Outro

More… if we have time

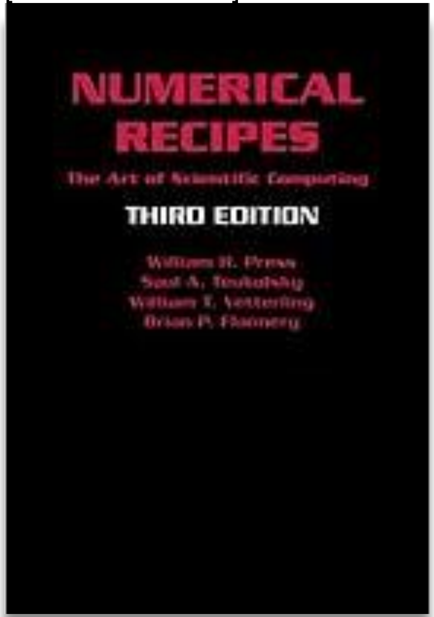
[G. Cowan+ ‘24]





Chap. 39 - Probability
Chap. 40 - Statistics

[W. H. Press+ ‘07]



Chap. 14 - Statistical Description of Data
Chap. 15 - Modeling of Data

Lesson adapted from my Master 2 notes, available at this link

Elements of statistics

Master 2 - Nuclei, Particles, Astroparticles and Cosmology

NPAC 2024–2025, Jonathan Biteau

This document aims to introduce and illustrate the essential elements at Master's level for performing statistical inference. The process of inference is understood here as evaluating the relevance of a model with respect to a set of data. A concise bibliography can be found at the end of this document to start exploring the statistical approaches employed in high-energy (astro)physics and cosmology. The content of this document is largely inspired by the chapters "*Probability*" [1] and "*Statistics*" [2] of the *Particle Data Group* review, whose reading is strongly encouraged. The same applies to the reference book *Numerical Recipes* [3], in particular chapters 14 "Statistical *Description of Data*" and 15 "*Modeling of Data*". The whole book *Numerical Recipes* can be considered an essential prerequisite for a thesis in our fields. The book provides examples in C/C++, which serve as an excellent guide to understanding the analytic approaches. The present document favours the use of Python libraries, which can be tested in the exercises.

I. Gamma-ray astronomy as a counting experiment

Concepts: *DL3-DL4-DL5, forward-folding, binomial / Poisson / Gauss, PMF / PDF / CDF / SF*

II. Statistical inference

Concepts: *Bayesian & frequentist, marginalizing & profiling, parameter estimation, (saturated) deviance*

III. Model selection

Concepts: *goodness of fit, residuals & pulls, under-/over-fitting,*
local & global p-values (trial factors / look-elsewhere), likelihood ratio test, significance level, Li & Ma

IV. Dealing with uncertainties

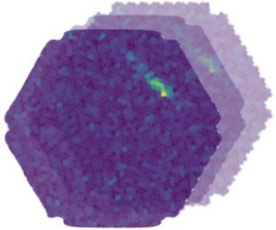
Concepts: *covariance, decorrelation energy, bowtie, systematic uncertainties, bracketing / marginalizing*

V. Outro

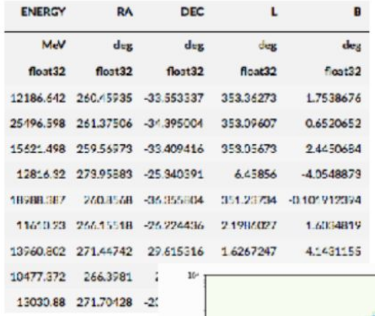
More… if we have time

Gamma-ray astronomy as a counting experiment

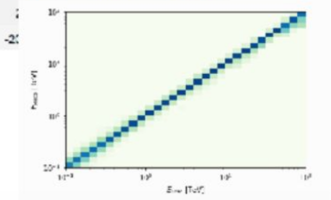
DL 0-1	DL 2	DL 3	DL 4	DL 5
raw - calibrated	*reconstructed*	*filtered γ-like events*	*science*	*high-level*



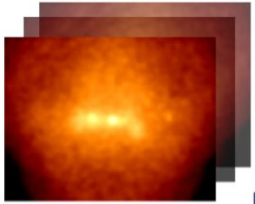
camera data



ENERGY	RA	DEC	L	B
MeV	deg	deg	deg	deg
float32	float32	float32	float32	float32
12186.642	260.45935	-33.553337	353.36273	1.7538676
25196.598	261.37506	-34.395004	353.09607	0.6520652
15621.498	259.56973	-33.409416	353.05673	2.4450684
12816.32	273.95883	-25.340391	6.45856	-4.0548873
[illegible]	[illegible]	[illegible]	[illegible]	[illegible]
[illegible]	[illegible]	-26.224436	[illegible]	[illegible]
13960.802	271.44742	29.615316	1.6267247	4.1431155
10477.372	266.3981	[illegible]	[illegible]	[illegible]
13030.88	271.70128	[illegible]	[illegible]	[illegible]



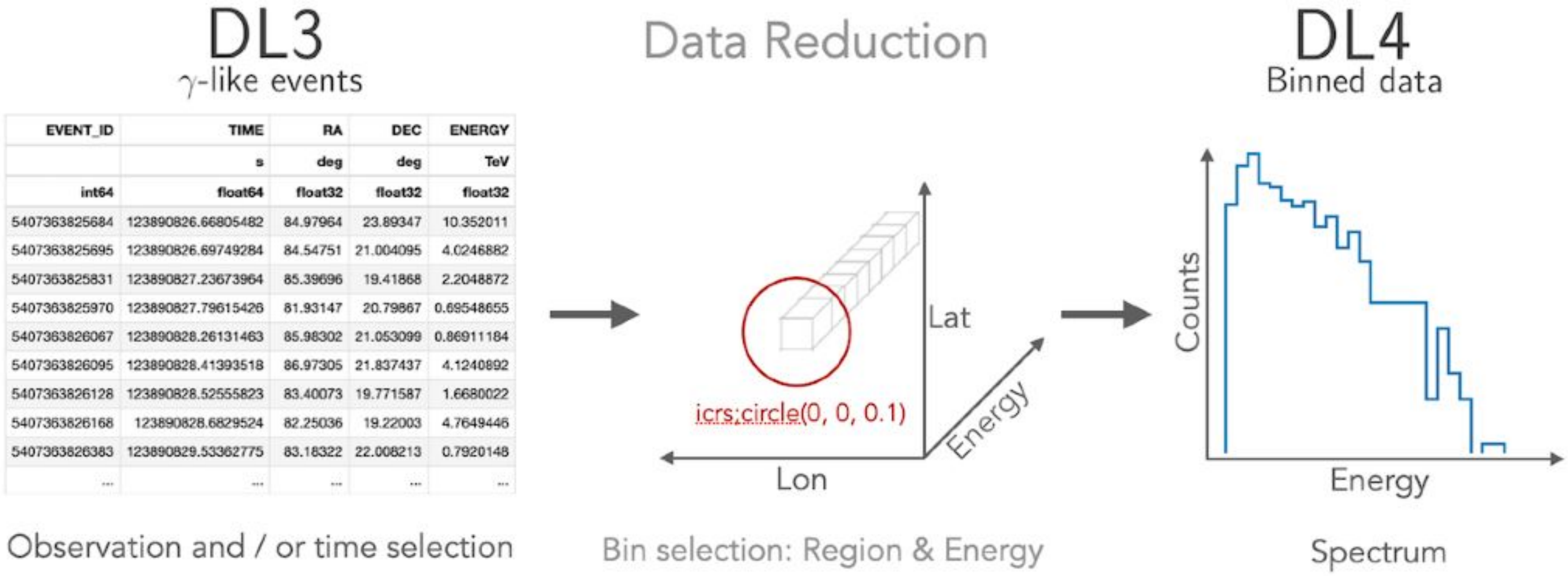
γ-like event lists
IRFs



maps,
spectra,
light curves

[GammaPy team, CTAO School 2024]

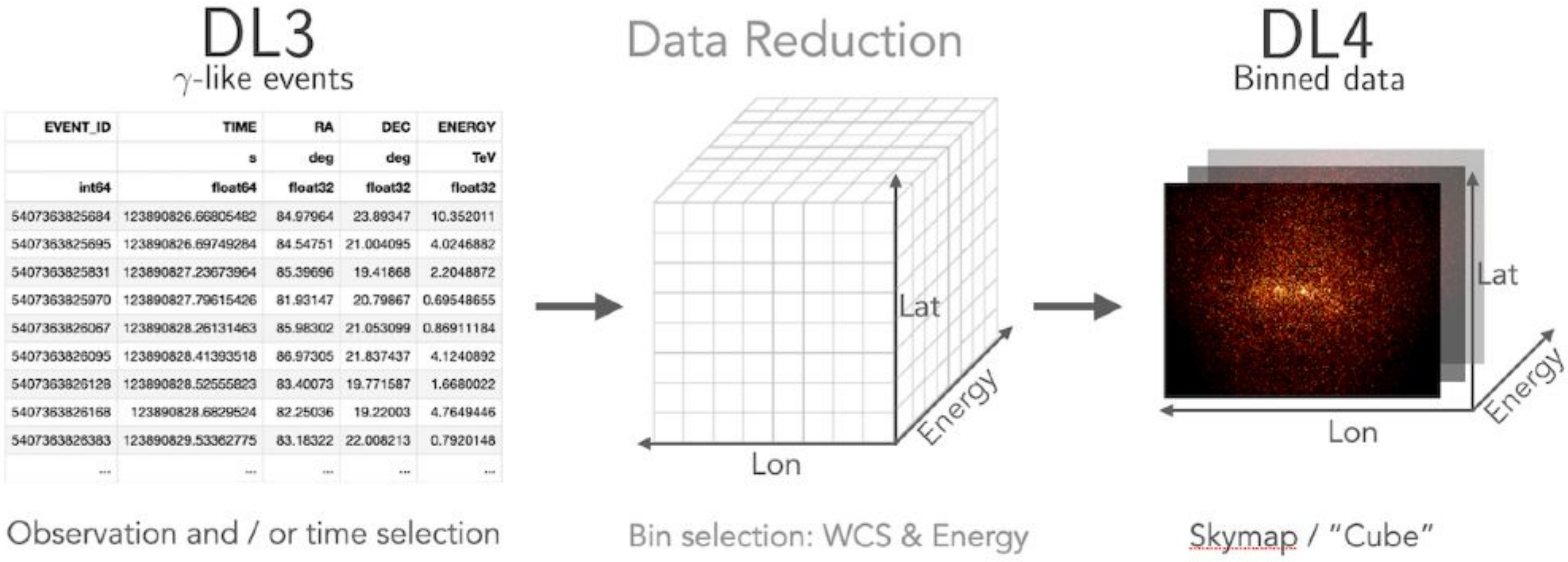
Gamma-ray astronomy as a counting experiment



EVENT_ID	TIME	RA	DEC	ENERGY
	s	deg	deg	TeV
int64	float64	float32	float32	float32
5407363825684	123890826.66805482	84.97964	23.89347	10.352011
5407363825695	123890826.69749284	84.54751	21.004095	4.0246882
5407363825831	123890827.23673964	85.39696	19.41868	2.2048872
5407363825970	123890827.79615426	81.93147	20.79867	0.69548655
5407363826067	123890828.26131463	85.98302	21.053099	0.86911184
5407363826095	123890828.41393518	86.97305	21.837437	4.1240892
5407363826128	123890828.52555823	83.40073	19.771587	1.6680022
5407363826168	123890828.6829524	82.25036	19.22003	4.7649446
5407363826383	123890829.53362775	83.18322	22.008213	0.7920148
...	...	...	...	...

[GammaPy team, CTAO School 2024]

Gamma-ray astronomy as a counting experiment



EVENT_ID	TIME	RA	DEC	ENERGY
	s	deg	deg	TeV
int64	float64	float32	float32	float32
5407363825684	123890826.66805482	84.97964	23.89347	10.352011
5407363825695	123890826.69749284	84.54751	21.004095	4.0246882
5407363825831	123890827.23673964	85.39696	19.41868	2.2048872
5407363825970	123890827.79615426	81.93147	20.79867	0.69548655
5407363826067	123890828.26131463	85.98302	21.053099	0.86911184
5407363826095	123890828.41393518	86.97305	21.837437	4.1240892
5407363826128	123890828.52555823	83.40073	19.771587	1.6680022
5407363826168	123890828.6829524	82.25036	19.22003	4.7649446
5407363826383	123890829.53362775	83.18322	22.008213	0.7920148
...	...	...	...	...

[GammaPy team, CTAO School 2024]

Gamma-ray astronomy as a counting experiment

2.2 IRF factorisation

Equation 2.2 implies 7-dimensional instrument response functions that in general are computationally unmanageable. Simplifications can be achieved by making further assumptions, and in existing Imaging Air Cherenkov Telescope (IACT) experiments the IRF is generally factorised as follows:

$$R_i(\hat{\alpha}, \hat{\delta}, \hat{E}|\alpha, \delta, E, t) = \mathrm{A}_i(\alpha, \delta, E, t) \times \mathrm{PSF}_i(\hat{\alpha}, \hat{\delta}|\alpha, \delta, E, t) \times \mathrm{D}_i(\hat{E}|\alpha, \delta, E, t) \quad (2.3)$$

where $\mathrm{A}_i(\alpha, \delta, E, t)$ is the effective area in units of cm^2, $\mathrm{PSF}_i(\hat{\alpha}, \hat{\delta}|\alpha, \delta, E, t)$ is the point spread function in units of sr^{-1}, with

$$\int d\hat{\Omega}\ \mathrm{PSF}_i(\hat{\alpha}, \hat{\delta}|\alpha, \delta, E, t) = 1 \quad (2.4)$$

and $\mathrm{D}_i(\hat{E}|\alpha, \delta, E, t)$ is the energy dispersion in units of TeV^{-1}, with

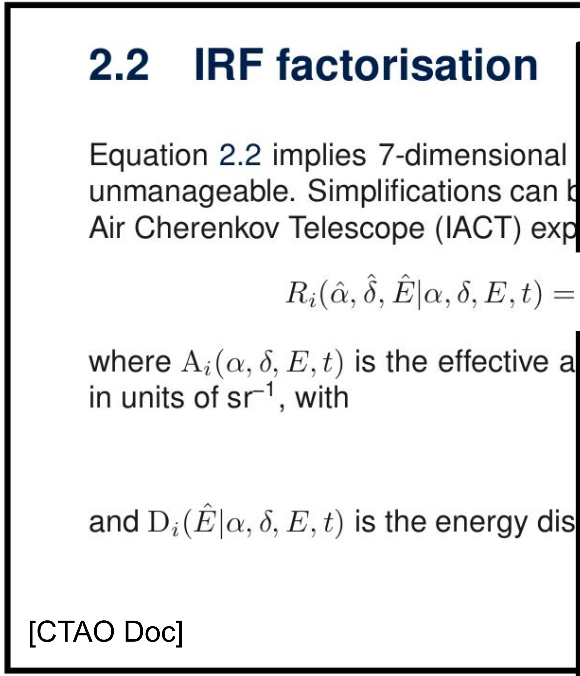
$$\int d\hat{E}\, \mathrm{D}_i(\hat{E}|\alpha, \delta, E, t) = 1 \quad (2.5)$$

[CTAO Doc]

Forward folding

Convolve true emission (i.e. model as a function of true energy, direction…)
with transfer matrix (Instrument Response Function linking true and observed quantities)
to enable a comparison to data in the observed space

Gamma-ray astronomy as a counting experiment



2.2 IRF factorisation

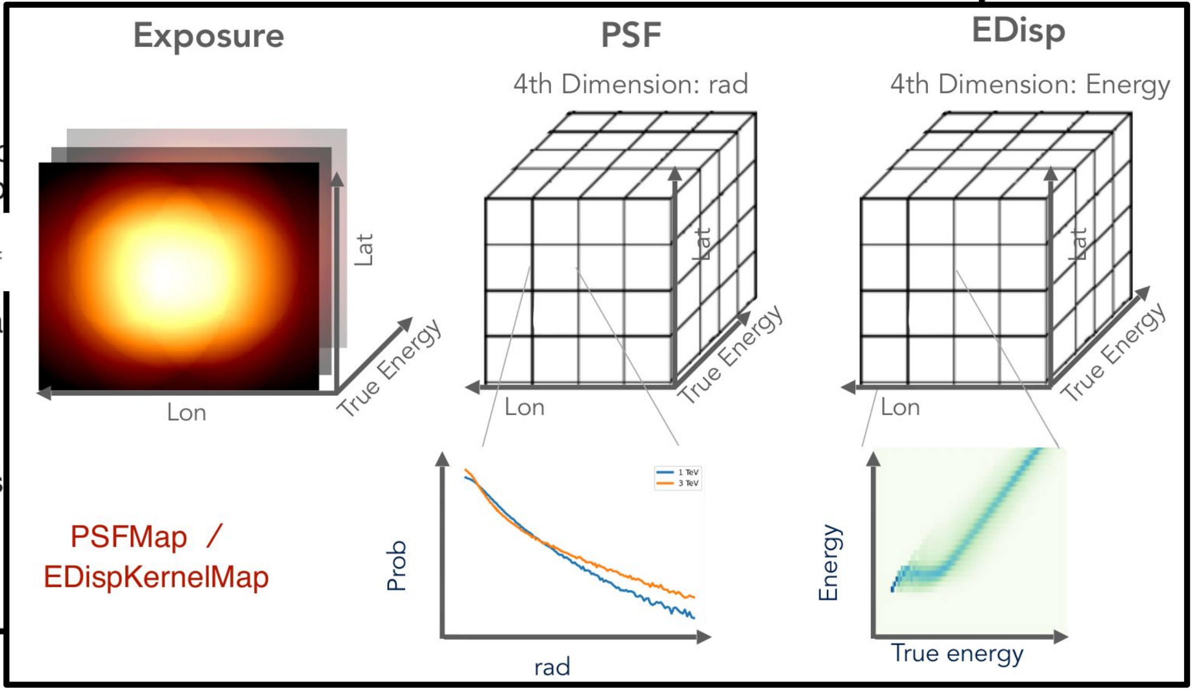
Equation 2.2 implies 7-dimensional
unmanageable. Simplifications can b
Air Cherenkov Telescope (IACT) exp

$$R_i(\hat{\alpha}, \hat{\delta}, \hat{E}|\alpha, \delta, E, t) =$$

where $\mathrm{A}_i(\alpha, \delta, E, t)$ is the effective a
in units of sr^{-1}, with

and $\mathrm{D}_i(\hat{E}|\alpha, \delta, E, t)$ is the energy dis

[CTAO Doc]



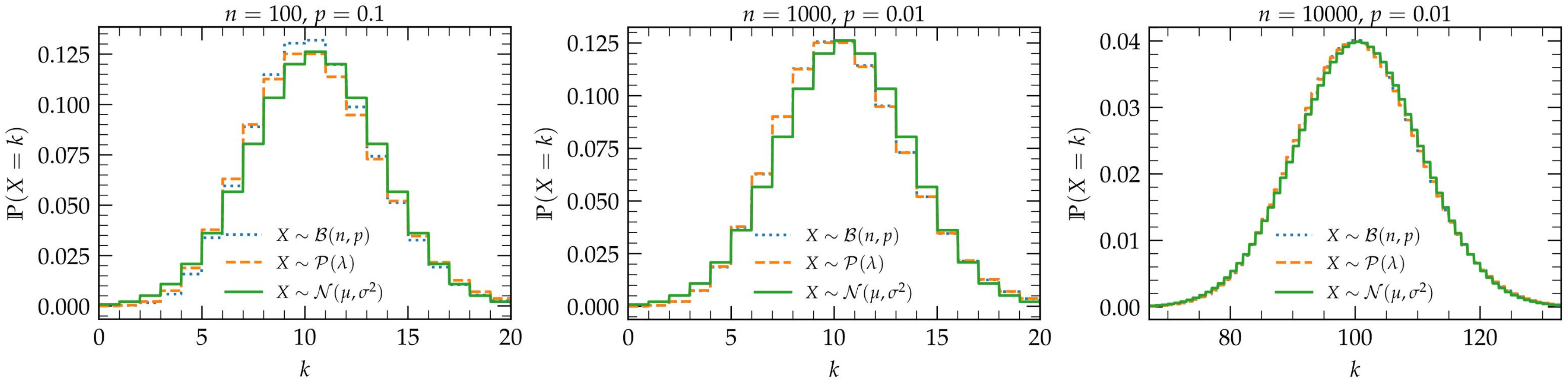
Forward folding

Convolve true emission (i.e. model as a function of true energy, direction…)
with transfer matrix (Instrument Response Function linking true and observed quantities)
to enable a comparison to data in the observed space

[GammaPy team, CTAO School 2024]

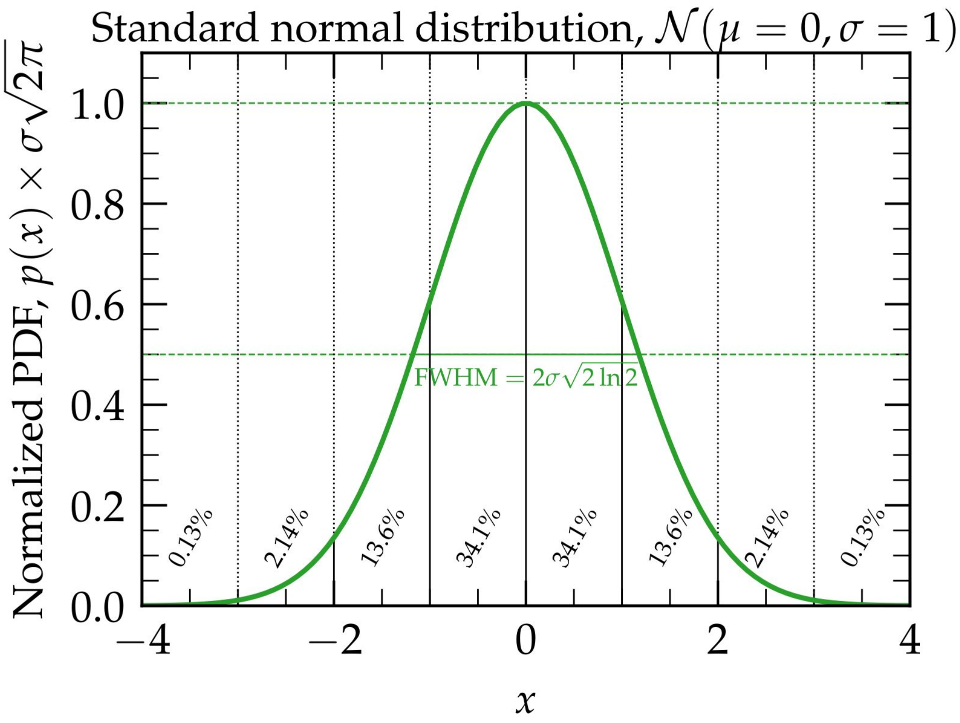
Gamma-ray astronomy as a counting experiment

When Bernoulli meets Poisson and Gauss



Gamma-ray astronomy as a counting experiment

From Gauss PMF to Gauss PDF



I. Gamma-ray astronomy as a counting experiment

Concepts: *DL3-DL4-DL5, forward-folding, binomial / Poisson / Gauss, PMF / PDF / CDF / SF*

II. Statistical inference

Concepts: *Bayesian & frequentist, marginalizing & profiling, parameter estimation, (saturated) deviance*

III. Model selection

Concepts: *goodness of fit, residuals & pulls, under-/over-fitting,*
local & global p-values (trial factors / look-elsewhere), likelihood ratio test, significance level, Li & Ma

IV. Dealing with uncertainties

Concepts: *covariance, decorrelation energy, bowtie, systematic uncertainties, bracketing / marginalizing*

V. Outro

More… if we have time

Statistical inference

Hessian and covariance matrix

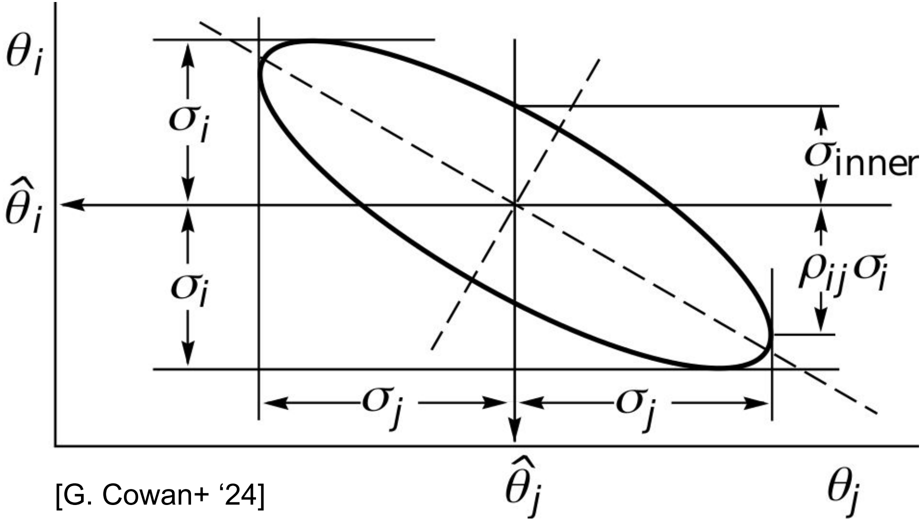
```
OptimizeResult

        backend    : minuit
        method     : migrad
        success    : True
        message    : Optimization terminated successfully.
        nfev       : 244
        total stat : 86.12

CovarianceResult

        backend    : minuit
        method     : hesse
        success    : True
        message    : Hesse terminated successfully.
```

[GammaPy team]



[G. Cowan+ ‘24]

I. Gamma-ray astronomy as a counting experiment

Concepts: *DL3-DL4-DL5, forward-folding, binomial / Poisson / Gauss, PMF / PDF / CDF / SF*

II. Statistical inference

Concepts: *Bayesian & frequentist, marginalizing & profiling, parameter estimation, (saturated) deviance*

III. Model selection

Concepts: *goodness of fit, residuals & pulls, under-/over-fitting, local & global p-values (trial factors / look-elsewhere), likelihood ratio test, significance level, Li & Ma*

IV. Dealing with uncertainties

Concepts: *covariance, decorrelation energy, bowtie, systematic uncertainties, bracketing / marginalizing*

V. Outro

More… if we have time

I. Gamma-ray astronomy as a counting experiment

Concepts: *DL3-DL4-DL5, forward-folding, binomial / Poisson / Gauss, PMF / PDF / CDF / SF*

II. Statistical inference

Concepts: *Bayesian & frequentist, marginalizing & profiling, parameter estimation, (saturated) deviance*

III. Model selection

Concepts: *goodness of fit, residuals & pulls, under-/over-fitting,*
local & global p-values (trial factors / look-elsewhere), likelihood ratio test, significance level, Li & Ma

IV. Dealing with uncertainties

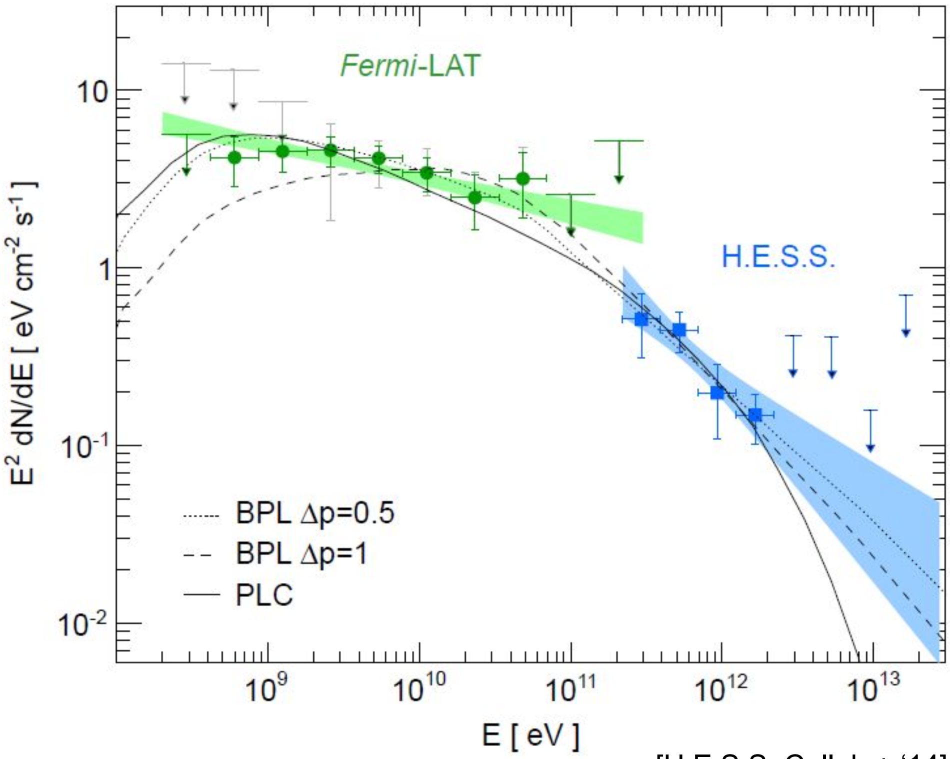
Concepts: *covariance, decorrelation energy, bowtie, systematic uncertainties, bracketing / marginalizing*

V. Outro

More… if we have time

Dealing with uncertainties

Bowtie / butterfly



[H.E.S.S. Collab.+ ‘14]

I. Gamma-ray astronomy as a counting experiment

Concepts: *DL3-DL4-DL5, forward-folding, binomial / Poisson / Gauss, PMF / PDF / CDF / SF*

II. Statistical inference

Concepts: *Bayesian & frequentist, marginalizing & profiling, parameter estimation, (saturated) deviance*

III. Model selection

Concepts: *goodness of fit, residuals & pulls, under-/over-fitting, local & global p-values (trial factors / look-elsewhere), likelihood ratio test, significance level, Li & Ma*

IV. Dealing with uncertainties

Concepts: *covariance, decorrelation energy, bowtie, systematic uncertainties, bracketing / marginalizing*

V. Outro

More… if we have time