$P(Q = v \mid I)$ test hypothesis

Assess what values of Q are consistent with data

$P(\Theta \mid I)$ where $\Theta = \underbrace{(\theta_1, \theta_2 \dots \theta_?)}_{\text{all possible values}}$

mass Higgs	$\mathbb{R}^+$	LHC data
position star in MW (x, y, z)	$\mathbb{R}^3$	parallax

Parameter Estimation

$$P(\Theta \mid d, I) = \frac{P(\Theta \mid I)\, P(d \mid \Theta, I)}{P(d \mid I)}$$

<u>evidence</u> : prob. of observing the data <u>at all</u>.

$$\underbrace{P(\Theta \mid d, I)}_{\text{posterior}} \propto \underbrace{P(\Theta \mid I)}_{\text{prior}}\, \underbrace{P(d \mid \Theta, I)}_{\text{likelihood}}$$

<u>posterior</u>: what are good values Θ given the data

<u>prior</u>: what are ~~plausi~~ possible values of Θ

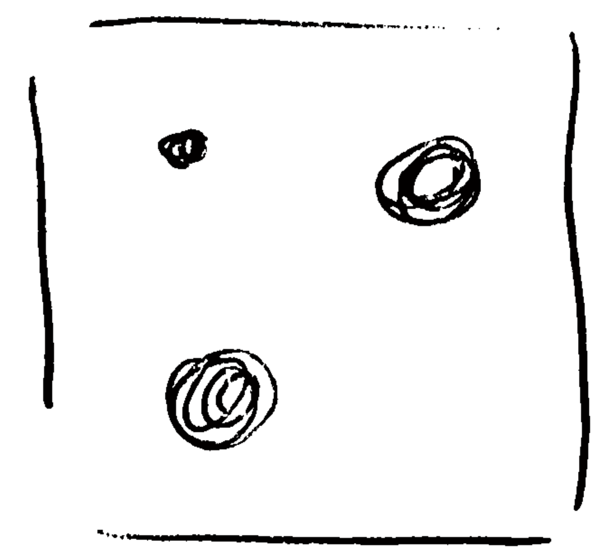
<u>likelihood</u>: describes the data generation

$$\vec{\theta} = (\underset{\substack{\uparrow \\ \text{target} \\ \text{(interesting)}}}{\vec{\theta}_1}, \underset{\substack{\uparrow \\ \text{nuisance}}}{\vec{\theta}_2})$$

$$P(\theta_1, \theta_2 | d, I) \propto \underline{P(\theta_1, \theta_2 | I)}\, P(d | \theta_1, \theta_2 I)$$

Marginalise over nuisance θ_2

$$P(\theta_1 | d, I) = \int d\theta_2 \; P(\theta_1, \theta_2 | d, I)$$



$L = \underline{\sigma_T} A T^4$

we need ~~to~~ account for variability of A, T

$\theta_1 : \sigma_T$

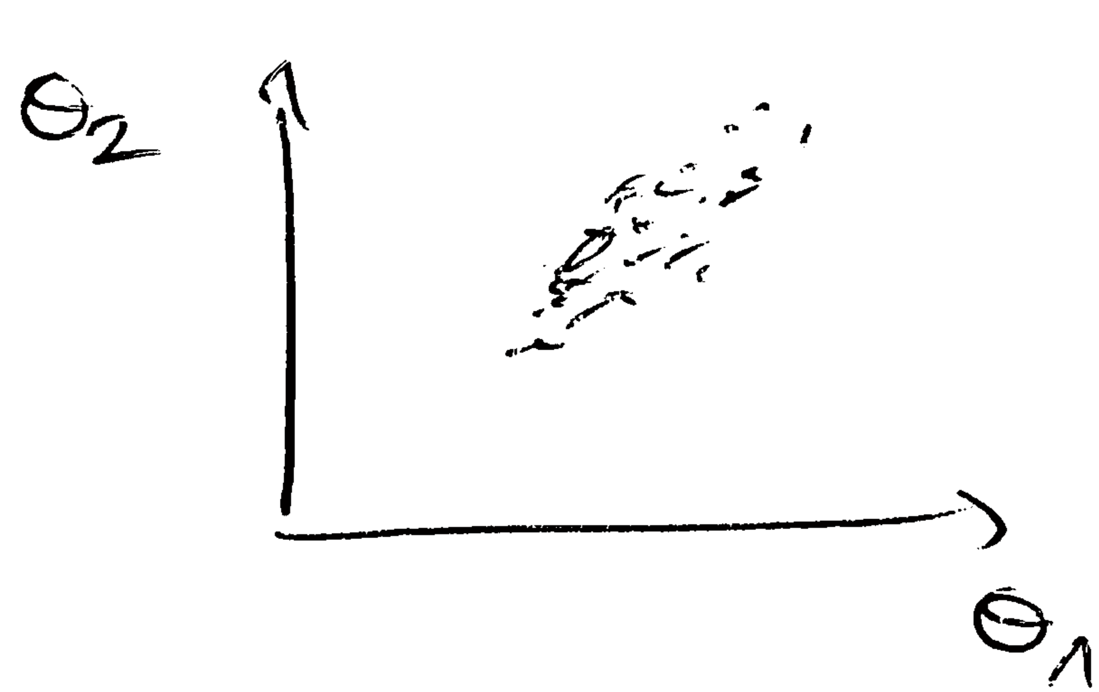
$\theta_2 : A, T$

θ_1 θ_2 are independent

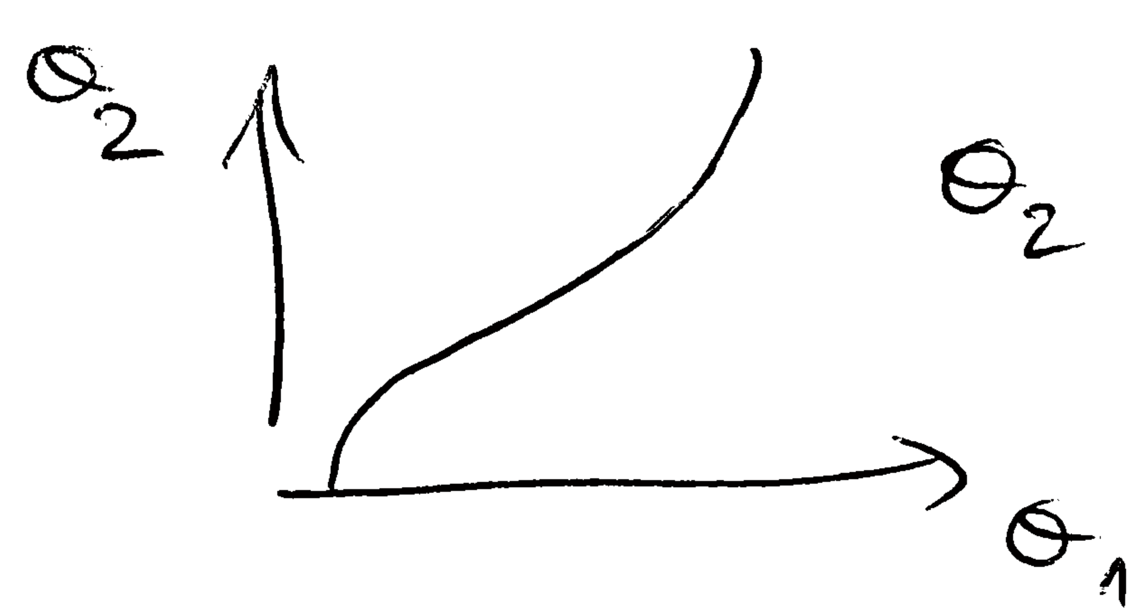
$$P(\theta_1, \theta_2 | I) = P(\theta_1 | I)\, P(\theta_2 | I)$$

θ_1 θ_2 are not independent

$$P(\theta_1, \theta_2 | I) = P(\theta_1 | I)\, P(\theta_2 | \theta_1 I)$$



$P(\theta_2 | \theta_1 I)$



$\theta_2 = f(\theta_1)$

$$P(\theta_2 | \theta_1, I) = \delta_D [\theta_2 - f(\theta_1)]$$

$\delta_D(x-m)$

physical laws:

$E = mc^2$

$\delta_D(E - mc^2)$

Recipe for Parameter Estimation

① What are the parameters? $\theta \begin{cases} \text{target} \\ \text{nuisance} \end{cases}$

② What prior information do we have?

$$I \longrightarrow P(\theta | I) \qquad \boxed{\text{prior}}$$

③ What is the data? $\vec{d}$

④ Make a probabilistic model of observation

$$\theta \longrightarrow d \qquad P(d | \theta) \quad \boxed{\text{likelihood}}$$

⑤ Unnormalised posterior

$$\tilde{P}(\theta | d, I) \doteq P(\theta | I) \, P(d | \theta)$$

⑥ Normalise posterior (check it is normalisable)

⑦ Marginalise over nuisance parameters

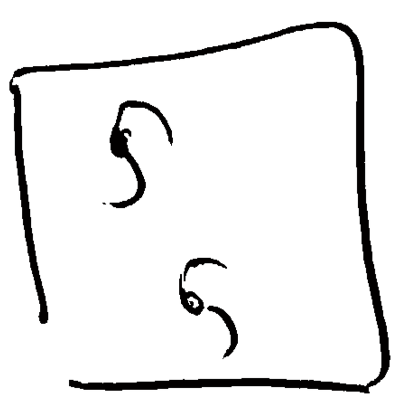
$$P(\theta_1 | d, I) = \int d\theta_2 \, P(\theta_1, \theta_2 | d, I)$$

⑧ Summaries ⟷ sampling.

best fitting value

(*) Watch out for wrong priors and model misspecification.

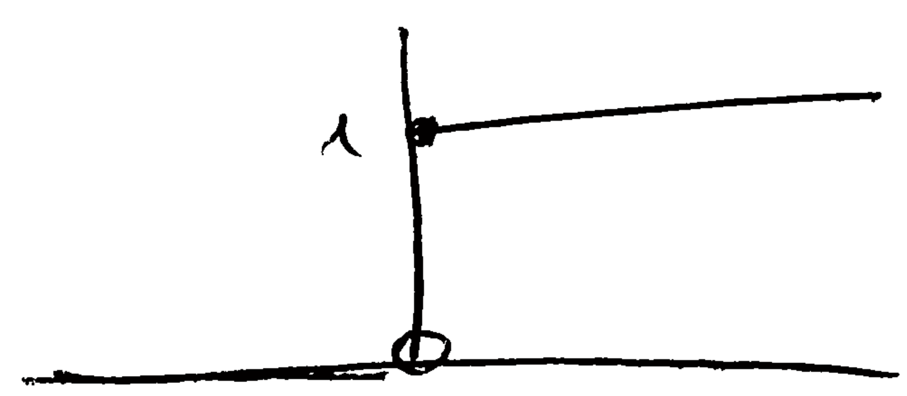
Example: Radio Sources



- radio sources $\Rightarrow$ flux S
- n sources observed $\rightarrow$ $S_1, S_2 \cdots S_n$
- number counts follow a power law

$$N(S) \propto \Theta(S)\, S^{-\alpha}$$

$\uparrow$ Heaviside funct.



$$\Theta \begin{cases} 0 & \text{if } x < 0 \\ 1 & \text{if } x > 0 \end{cases}$$

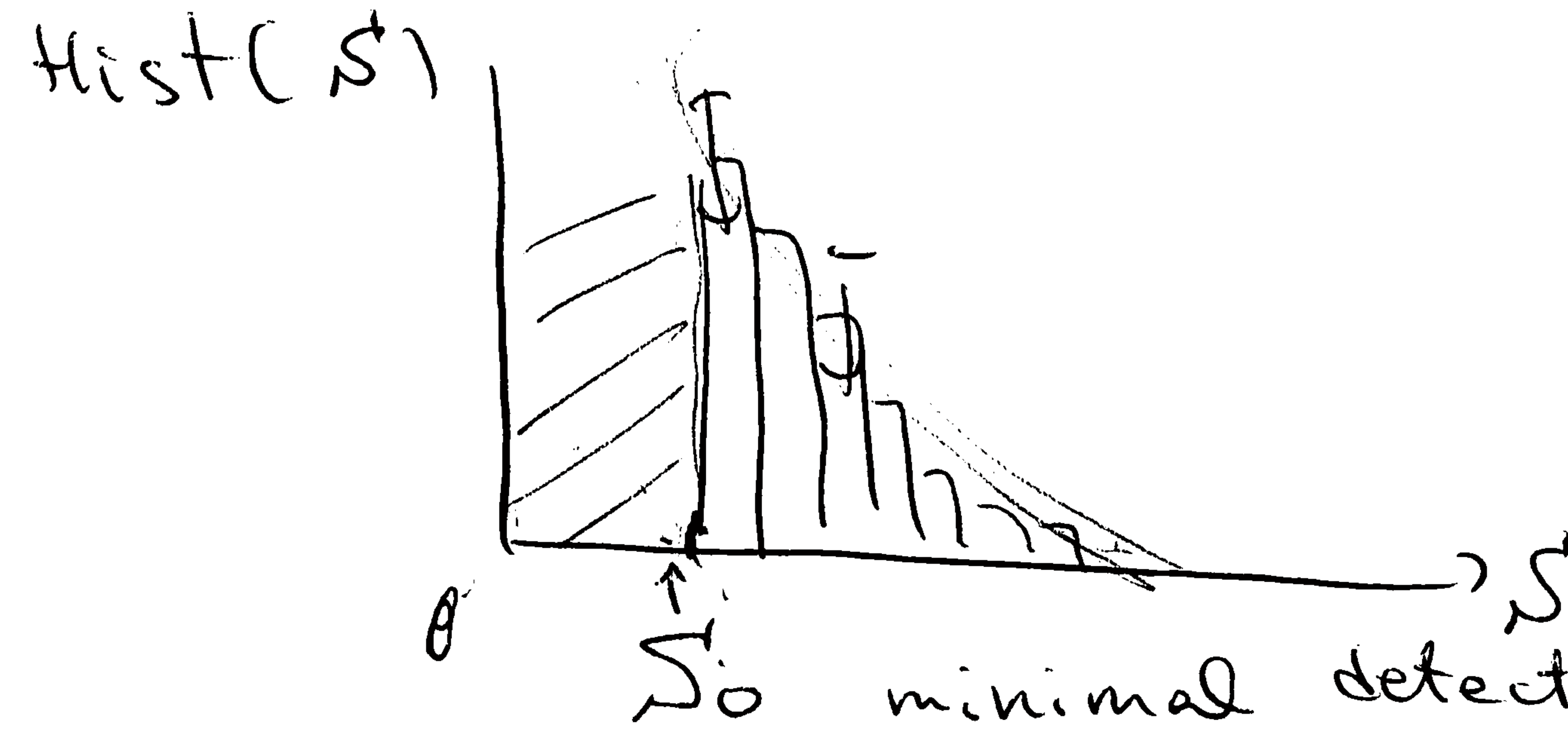
① parameter? α

② prior info?

$\alpha \geq 0$

$P(\alpha \mid I)$ is "smooth"

③ data? $\vec{S} = S_{1:n} = S_1, S_2 \cdots S_n$



④ data model

$$\underbrace{P(\underbrace{S_{1:n}}_{\text{data}} \mid \underbrace{\alpha}_{\text{param}}, \underbrace{n, S_0}_{\text{info.}})}_{\text{prob. of observing the full dataset.}} \underset{\text{independent}}{=} \prod_{i=1}^{n} \underbrace{P(S_i \mid \alpha, S_0)}_{\text{prob. of observing each source.}}$$

$$P(S_i \mid \alpha, S_0) \propto \underbrace{\Theta(S_i)\, S_i^{-\alpha}}_{\text{pow law}} \; \underbrace{\Theta(S_i - S_0)}_{\substack{\text{brighter than } S_0 \\ \text{(measurement)}}}$$

$$\propto \Theta(S_i - S_0)\, S_i^{-\alpha} \qquad \boxed{\text{likelihood}}$$

$S_0 > 0$

5 Normalisable?

$$N = \int_{S_0}^{\infty} dS_i \; S_i^{-\alpha} = \frac{1}{\alpha - 1} S_0^{-\alpha+1}$$

$\boxed{\alpha = 1}$ problem!

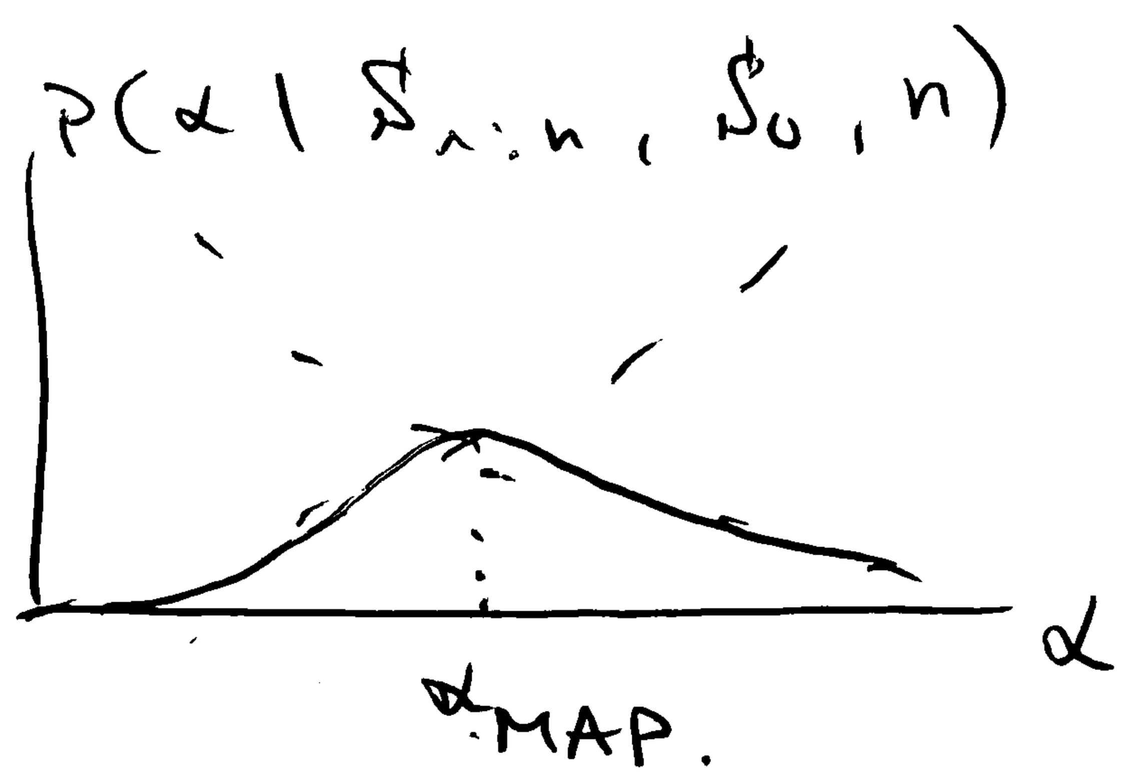
$$P(S_i \mid \alpha, S_0) = \Theta(S_i - S_0)\, \Theta(\alpha - 1)\, \frac{\alpha - 1}{S_0} \left(\frac{S_i}{S_0}\right)^{-\alpha}$$

$$P(S_{1:n} \mid \alpha, S_0, n) = \frac{(\alpha-1)^n}{S_0^n}\, \Theta(\alpha - 1) \prod_{i=1}^{n} \Theta(S_i - S_0) \left(\frac{S_i}{S_0}\right)^{-\alpha}$$

6 posterior?

$$P(\alpha \mid S_{1:n}, S_0, n) \propto$$

$$\propto \underbrace{P(\alpha \mid I)}_{\cancel{\Theta(\alpha)}}\, \Theta(\alpha - 1)\, \frac{(\alpha-1)^n}{S_0^n} \prod_{i=1}^{n} \left(\frac{S_i}{S_0}\right)^{-\alpha}$$



7 summaries

α_{MAP}