BASICS OF BAYESIAN DATA ANALYSIS

$$P(a,b) = P(a)\,P(b|a) \Rightarrow \boxed{p(b|a) = \frac{p(a|b)\,p(b)}{p(a)}}$$

BAYES THEOREM

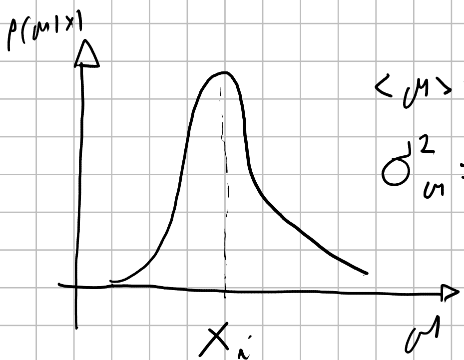
- $p(a|b)$: likelihood to have a given b
- $p(b)$: prior on b
- $p(b|a)$: posterior on b given our observation of a
- $P(a) = \int p(a|b)\,p(b)\,db \neq 1$ in general. it is evidence.

Example We want to estimate the μ of a Gaussian distribution.

$$p(x|\mu) = \frac{1}{\sqrt{2\pi}\,\sigma}\, e^{-\frac{(x-\mu)^2}{2\sigma^2}}$$

If we observe a value x_i $\Rightarrow$ $p(\mu|x_i) = \dfrac{p(x_i|\mu)\,\overbrace{p(\mu)}^{\text{UNIFORM}}}{\int p(x_i|\mu)\,p(\mu)\,d\mu}$

$$p(\mu|x_i) = \frac{1}{\sqrt{2\pi}\,\sigma}\, e^{-\frac{(x_i-\mu)^2}{2\sigma^2}}\, \frac{p(\mu)}{p(\mu)} \cdot \frac{1}{\underbrace{\int p(x_i|\mu)\,d\mu}_{1 \text{ FOR SYMMETRY}}} = \frac{1}{\sqrt{2\pi}\,\sigma}\, e^{-\frac{(x_i-\mu)^2}{2\sigma^2}}$$



$$\langle \mu \rangle = \int \mu\, p(\mu|x_i)\,d\mu = x_i$$
$$\sigma_\mu^2 = \int p(\mu|x_i)\,(\mu - \langle \mu \rangle)^2\,d\mu = \sigma^2$$

If we have multiple samples $\{X\} = \{x_1, \dots, x_N\}$

$$p(\{X\}|\mu) = \prod_{i=1}^{N} p(x_i|\mu) = \prod_{i=1}^{N} \frac{1}{\sqrt{2\pi}\,\sigma}\, e^{-\frac{(x_i-\mu)^2}{2\sigma^2}} = \left[\frac{1}{\sqrt{2\pi}\,\sigma}\right]^N \exp\left[\sum_{i=1}^{N} \frac{(x_i-\mu)^2}{2\sigma^2}\right]$$

$$\sum_{i=1}^{N} -\frac{(x_i-\mu)^2}{2\sigma^2} = -\sum \frac{x_i^2 - 2x_i\mu + \mu^2}{2\sigma^2} = -\frac{(\bar{x}-\mu)^2}{2\bar{\sigma}^2}$$

$$\bar{x} = \sum \frac{x_i}{N} \qquad \bar{\sigma}^2 = \frac{\sigma^2}{N}$$

$$P(\mu|\{x\}) = \frac{1}{\sqrt{2\pi}\,\bar{\sigma}} e^{-\frac{(\bar{x}-\mu)^2}{2\bar{\sigma}^2}}$$

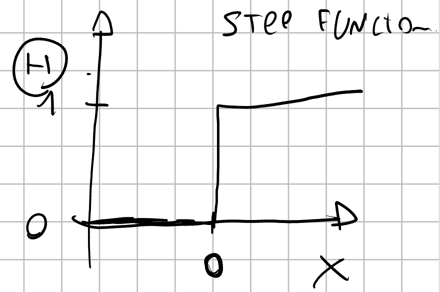
FROM HERE the $\sqrt{N}$ scaling of errors for the central limit theorem.

SELECTION BIAS

Let us assume that we are not able to observe all the x_i generated by the gaussian. Our experiment can only detect $x_i > 0$

$$\Rightarrow \mathcal{L}_D(X|\mu) = \frac{N(x,\mu)\,\Theta(x>0)}{\int N(x,\mu)\,\Theta(x>0)\,dx}$$

→ NEW NORMALIZATION



NOTE: Θ acts as some sort of detection probability of X. it is one when you detect x and $\emptyset$ otherwise

$\Theta \equiv P(DET|X)$

The new likelihood is therefore $\mathcal{L}_D(X|\mu) = \frac{\mathcal{L}(x|\mu)\,P(DET|X)}{\int_0^{+\infty}\mathcal{L}(x|\mu)\,dx}$

Now let's see the notebook N°1. When you collect x_i samples from your experiment.

$$\mathcal{L}_D(x|\mu) = \frac{\mathcal{L}(x|\mu)}{\int_0^{+\infty}\mathcal{L}(x|\mu)\,dx}$$

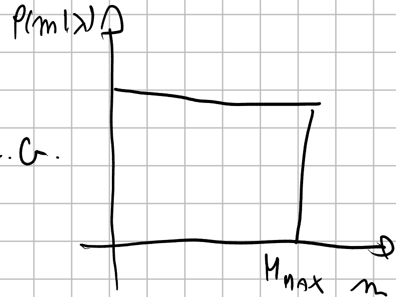
NOTE $P(DET|X_i) = 1$ BC WE DETECT IT.

GW POPULATION INFERENCE

FROM MANDEL ET AL MNRAS 2018

We want to estimate a population parameter λ in common to the entire GW population. For example λ can be the maximum mass of the BBH distribution

$p_{pop}(\theta | \lambda) \rightarrow$ POPULATION DISTRIBUTION e.g.



$\theta \rightarrow$ PARAMETERS OF THE GW. If we are perfectly able to detect all GW events and measure their parameters.

$$\theta_i \Rightarrow p(\{\theta\}|\lambda) = \prod_{i=1}^{N} \frac{p_{pop}(\theta_i|\lambda)}{\underbrace{\int d\theta \, p_{pop}(\theta|\lambda)}_{\rightarrow \text{ NORMALIZED USUALLY}}}$$

However our detectors have a selection bias (see example of a Gaussian).

$$\Rightarrow p_D(\{\theta\}|\lambda) = \prod_{i=1}^{N} \frac{p_{pop}(\theta_i|\lambda) \overbrace{p(DET|\theta_i)}^{1 \text{ AS WE DETECT THE EVENT}}}{\int d\theta \, p_{pop}(\theta|\lambda) \, p(DET|\theta)}$$

NOTE This is not like the gaussian as in GW detectors we have noise that it is fluctuating. I.e. our detection possibility is not 1 for a fixed θ

$$p(DET|\theta) = \int_{X \in DET} \underbrace{\mathcal{L}(X|\theta)}_{\rightarrow \text{ GW LIKELIHOOD}} d\theta$$

Moreover, we need to remember that you do not measure perfectly θ_i, e.g.

$$P(x|\lambda) = \int d\theta\, \mathcal{L}(x|\theta)\, P(\theta|\lambda) = \prod_{i=1}^{N} \frac{\int d\theta\, \mathcal{L}(x|\theta)\, P_{pop}(\theta|\lambda)}{\int d\theta\, P(DET|\theta)\, P_{pop}(\theta|\lambda)}$$

The posterior on the population parameters is therefore

PARAMETER ESTIMATION

$$P(\lambda|\{x\}) \propto P(\lambda) \prod_{i=1}^{N} \frac{\int d\theta\, \mathcal{L}(x|\theta)\, P_{pop}(\theta|\lambda)\, d\theta}{\int d\theta\, p(DET|\theta)\, P_{pop}(\theta|\lambda)\, d\theta}$$

Prior

SELECTION BIAS

An example with H_0

We want to measure H_0 with the EM counterpart technique.

$\lambda = H_0 \qquad \theta = z$

$$P(H_0|\{GRB\},\{GW\}) = P(H_0) \prod_{i=1}^{N} \frac{\int \mathcal{L}(GRB_i, GW_i|z)\, P_{pop}(z)\, dz}{\int P(DET_{GW}, DET_{GRB}|z)\, P(z)\, dz}$$

POSTERIOR ON z

$$\mathcal{L}(GRB_i, GW_i|z) = \underbrace{\mathcal{L}(GRB_i|z)}_{\delta(z-z_{EM})}\, \mathcal{L}(GW_i|z) = \delta(z-z_{EM})\, \frac{P(z|GW)}{\pi(z)}$$

PRIOR USED FOR PE

But GW does not give z, they give us d_L, so we need to do a change of variable $z \to d_L$.

PROB RULES

$$\mathcal{L}(GRB_i, GW_i|z) = \delta(z-z_{EM})\, \frac{P(d_L(z,H_0)|GW)\left|\frac{\partial d_L}{\partial z}\right|}{\pi(d_L(z,H_0)|GW)\left|\frac{\partial d_L}{\partial z}\right|}$$

$$\Rightarrow \mathcal{L}(GRB_i, GW_i | z) = \delta(z - z_{EM}) \frac{p(d_L(z, H_0) | GW)}{\pi(d_L(z, H_0) GW)}$$

We can do the integral in z and obtain

$$p(H_0 | DATA) = p(H_0) \prod_{i=1}^{N} \frac{p(d_L(z_{EM}, H_0) | GW)}{\pi(d_L(z_{EM}, H_0) GW)} \frac{1}{\int p(DET | z) p(z) dz}$$

What about the selection Bias? We will assume that

$$p(z) = \frac{z^2}{3 z^3_{MAX}}$$ and $$p(DET | z) = p(DET_{GW} | z) = \Theta(d_L(z, H_0) < D_H)$$

Namely I detect all GWs below a certain d_L

$$\int p(DET | z) p(z) = \int_0^{D_H \cdot H_0} \frac{z^2}{3 z^3_{MAX}} = \underbrace{\frac{D_H^3}{9 z^3_{MAX}}}_{\text{COSTANT, NOT IMPORTANT}} H_0^3 \propto H_0^3$$

. The selection Bias is proportional to H_0^3

$$\Rightarrow \boxed{p(H_0 | DATA) = p(H_0) \frac{1}{H_0^{3N}} \prod_{i=1}^{N} \frac{p(d_L(z_{EM}, H_0) | x)}{\pi(d_L(z_{EM}, H_0 | x)}}$$