Data

$y_1 \quad y_2 \quad \dots \quad y_n$

$x_1 \quad x_2 \quad \dots \quad x_n$

Model

$$\begin{cases} \text{for } i \text{ in } 1 \dots n \\ \quad y_i \sim N(\mu_i, \sigma^2) \\ \quad \mu_i = \beta_0 + \beta_1 x_i \end{cases}$$

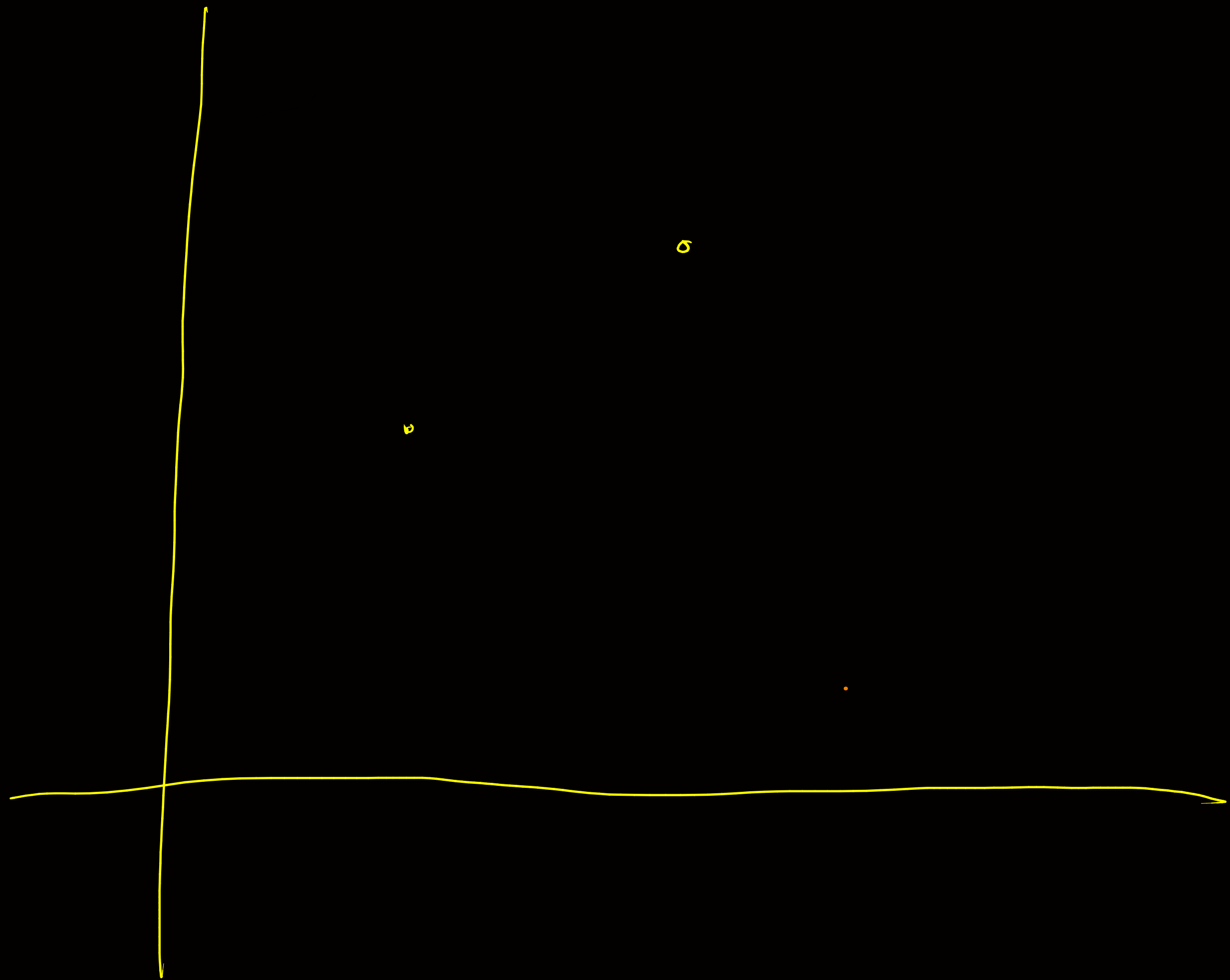
assumed model
normal
linear
model

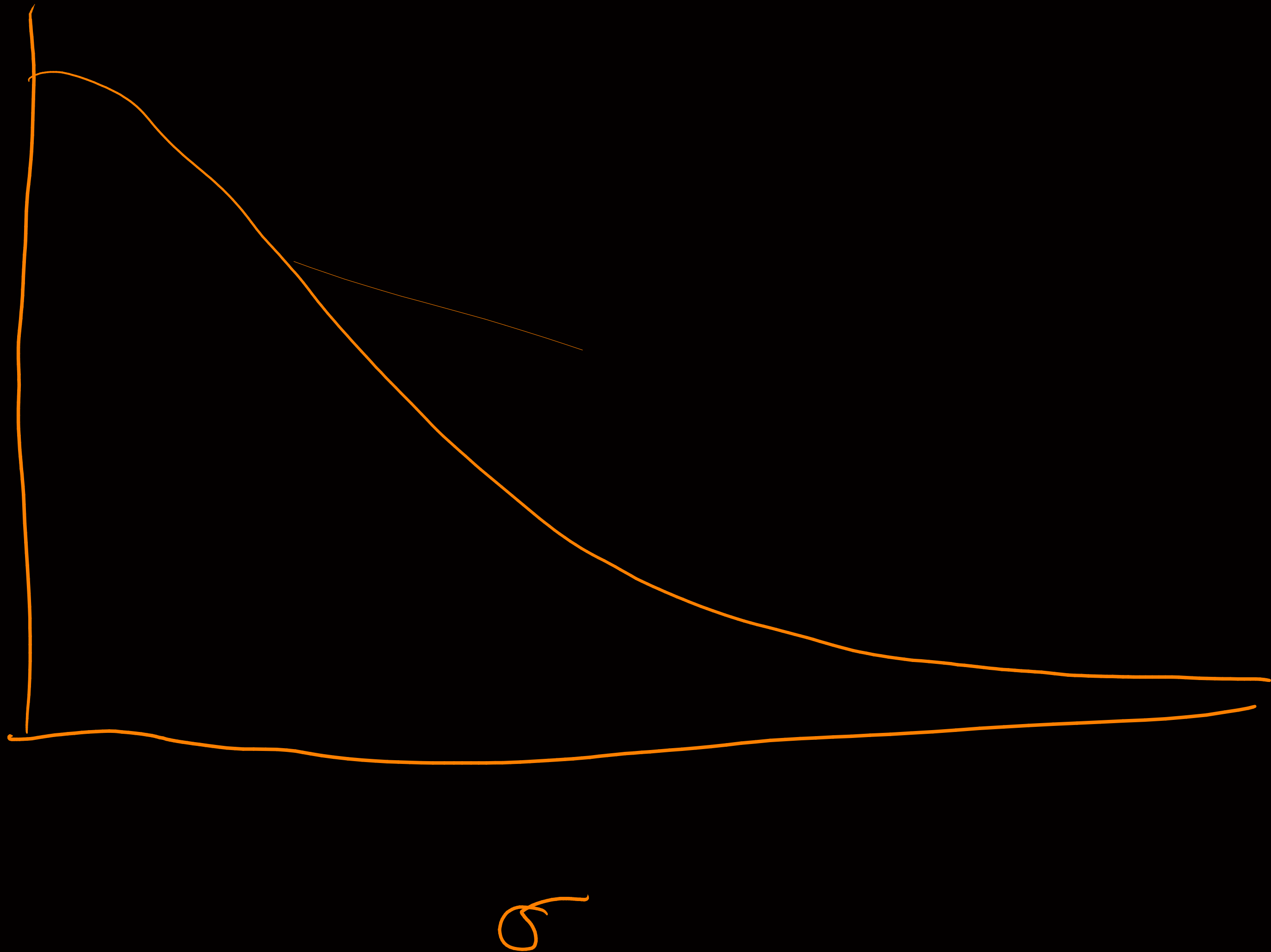
unknowns : $\beta_0, \beta_1, \sigma^2$

Inference

classical

Bayesian




σ

data model unknowns

Bayes's theorem

$$\underbrace{p(\text{unkowns} \mid \text{data}, \text{model})}_{\text{posterior}}$$

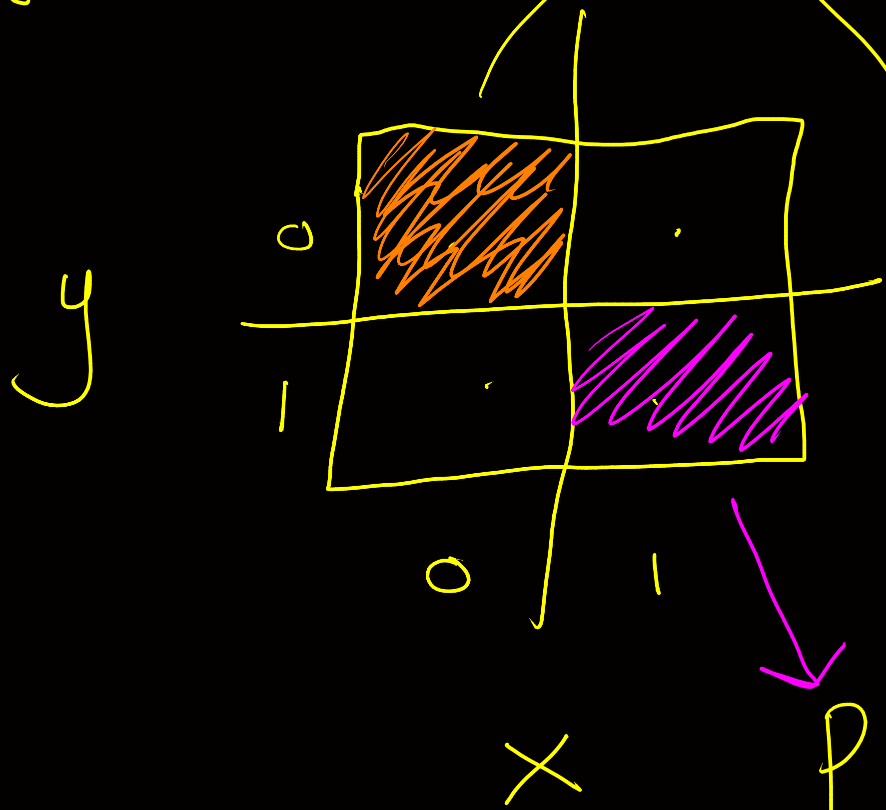
$\tilde{\beta}_0 \quad \tilde{\beta}_1 \quad \tilde{\sigma}$

$$p(\text{unknowns} = \tilde{\beta}_0, \tilde{\beta}_1, \tilde{\sigma} \mid \text{data}, \text{model})$$

$$= \frac{p(\text{data} \mid \tilde{\beta}_0, \tilde{\beta}_1, \tilde{\sigma}, \text{model})\, p(\tilde{\beta}_0, \tilde{\beta}_1, \sigma^2 \mid \text{model})}{\int p(\text{data} \mid \tilde{\beta}_0, \tilde{\beta}_1, \tilde{\sigma}\ \text{model})\, p(\tilde{\beta}, \tilde{\beta}, \sigma^2 \mid \text{model})}$$

$X \in \{0, 1\}$

$Y \in \{0, 1\}$



Joint prob

$p(x=0, y=0)$

$p(x=1, y=1)$

$$p(x=0, y=1)$$

$$= p(x=0 \mid y=1)\, p(y=1)$$

$$= p(y=1 \mid x=0)\, p(x=0)$$

$$p(y=1 \mid x=0)\, p(x=0) = p(x=0 \mid y=1)\, p(y=1)$$

$$p(y=1 \mid x=0) = \frac{p(x=0 \mid y=1)\, p(y=1)}{p(x=0)}$$

$y_1, y_2, \dots, y_n$ $\quad n = 250$

Prob of Heads

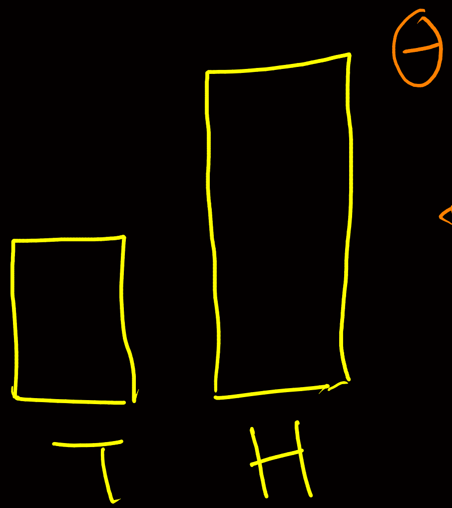
$y_i \in \{H, T\}$

{

for i in $1 \dots n$

$y_i \sim \text{Bernoulli}(\theta)$

}



Bernoulli

$y_1 \; y_2 \; \cdots \; y_n$

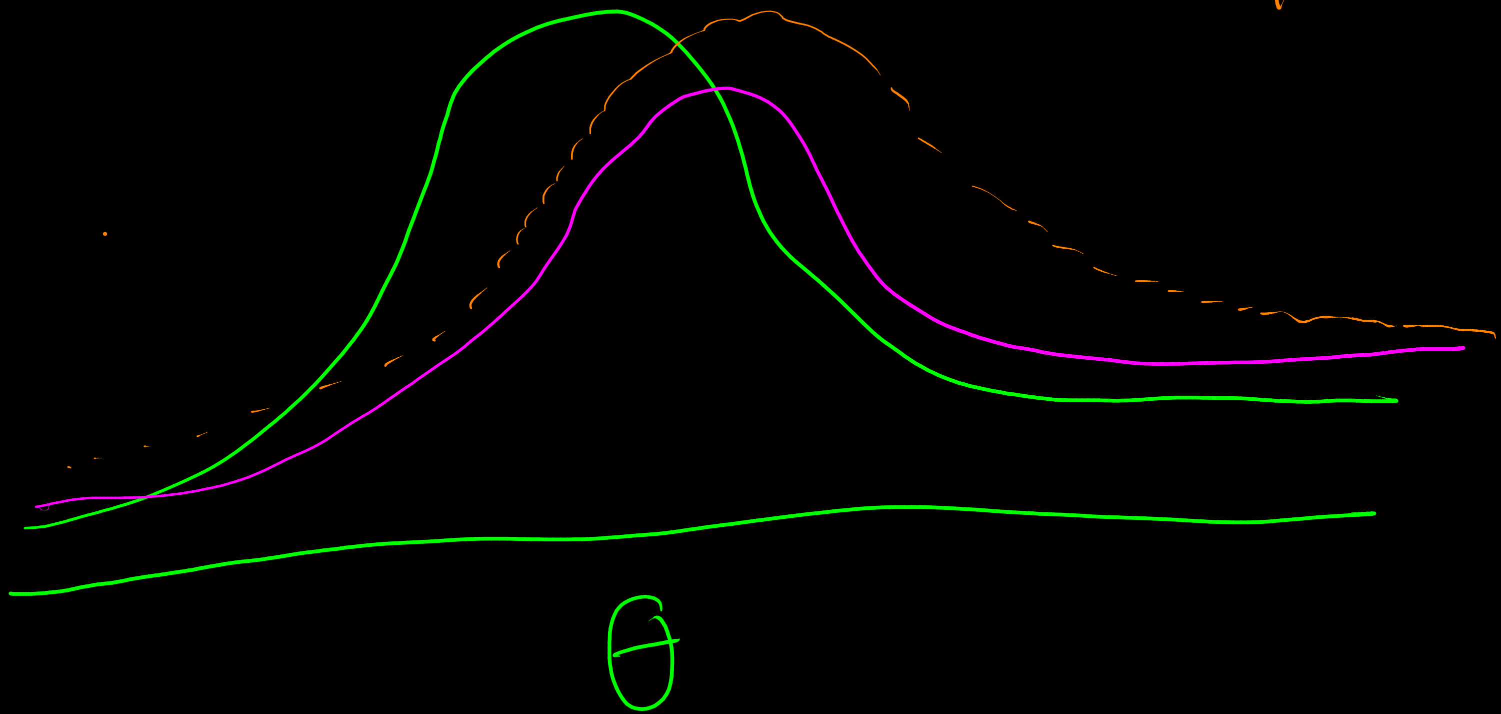
θ

Likelihood : $p(y_1 \ldots y_n | \theta)$

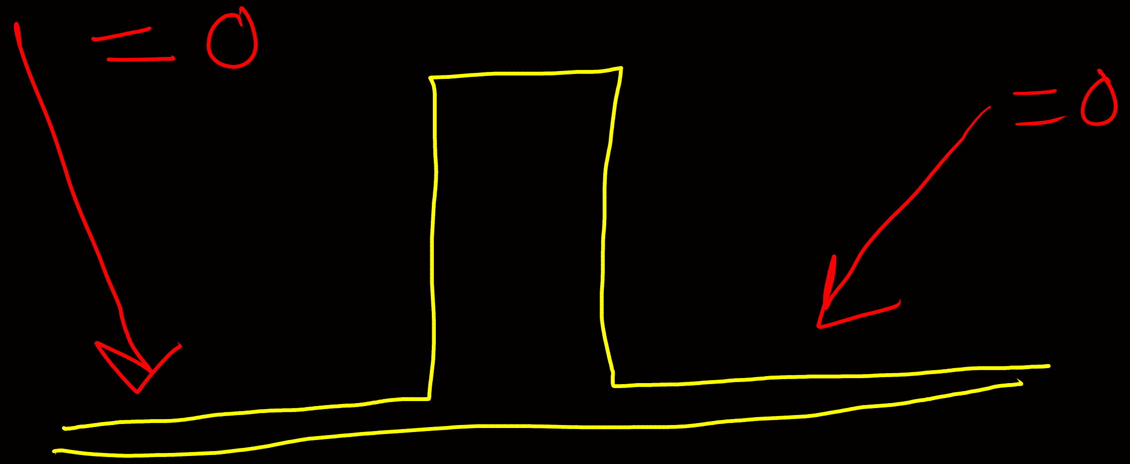
$$= \prod p(y_i | \theta)$$

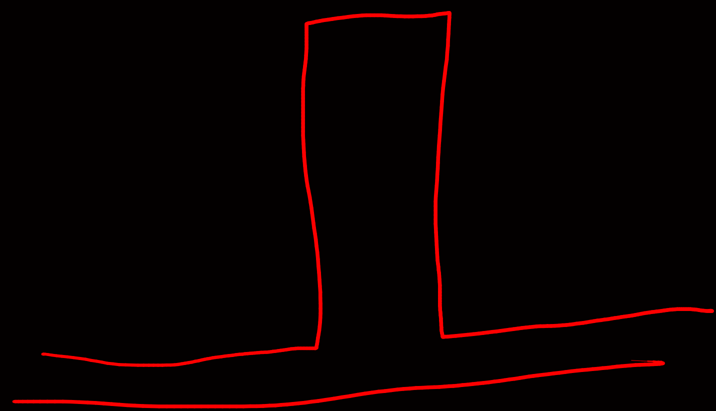
$$\frac{p(D | \theta = 0.4)}{p(D | \theta = 0.6)}$$

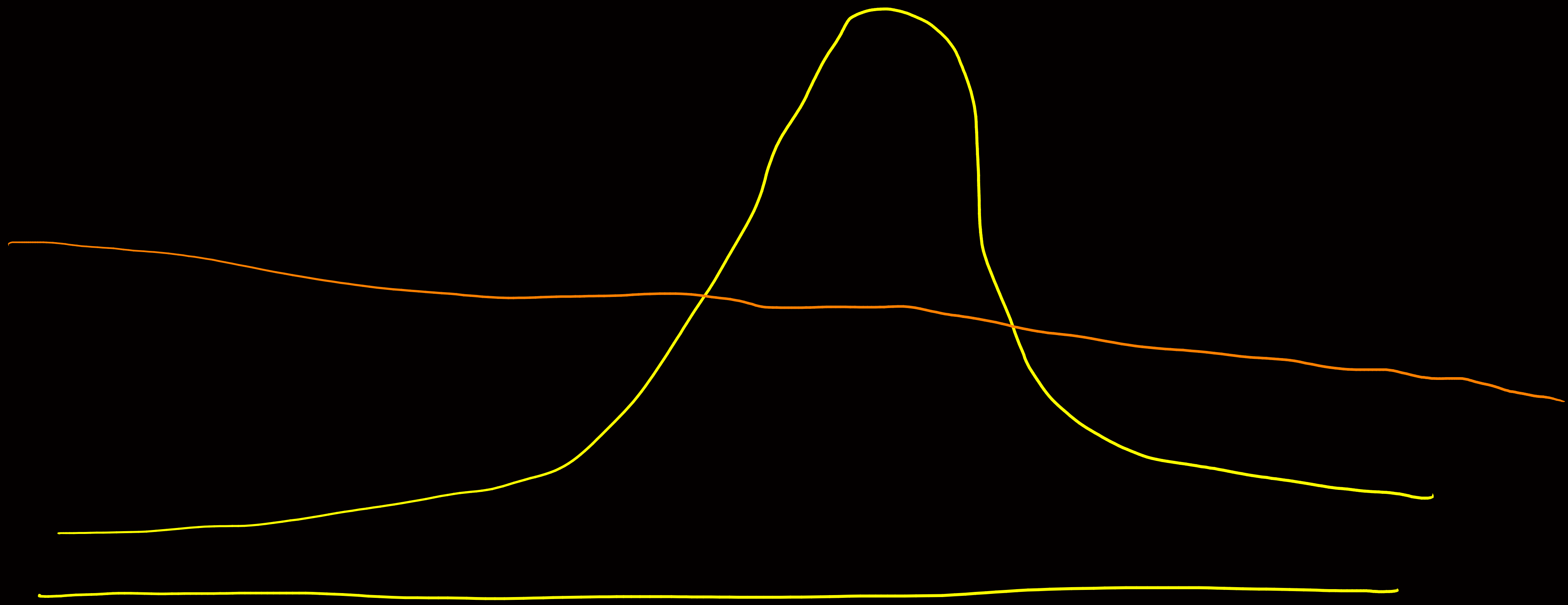
$$p(\theta|D) = \frac{p(D|\theta)\,p(\theta)}{\underbrace{\int_0^1 p(D|\theta)\,p(\theta)\,d\theta}_{\text{marginal likelihood}}}$$



$$p(\theta|D) \propto p(D|\theta) \times p(\theta)$$







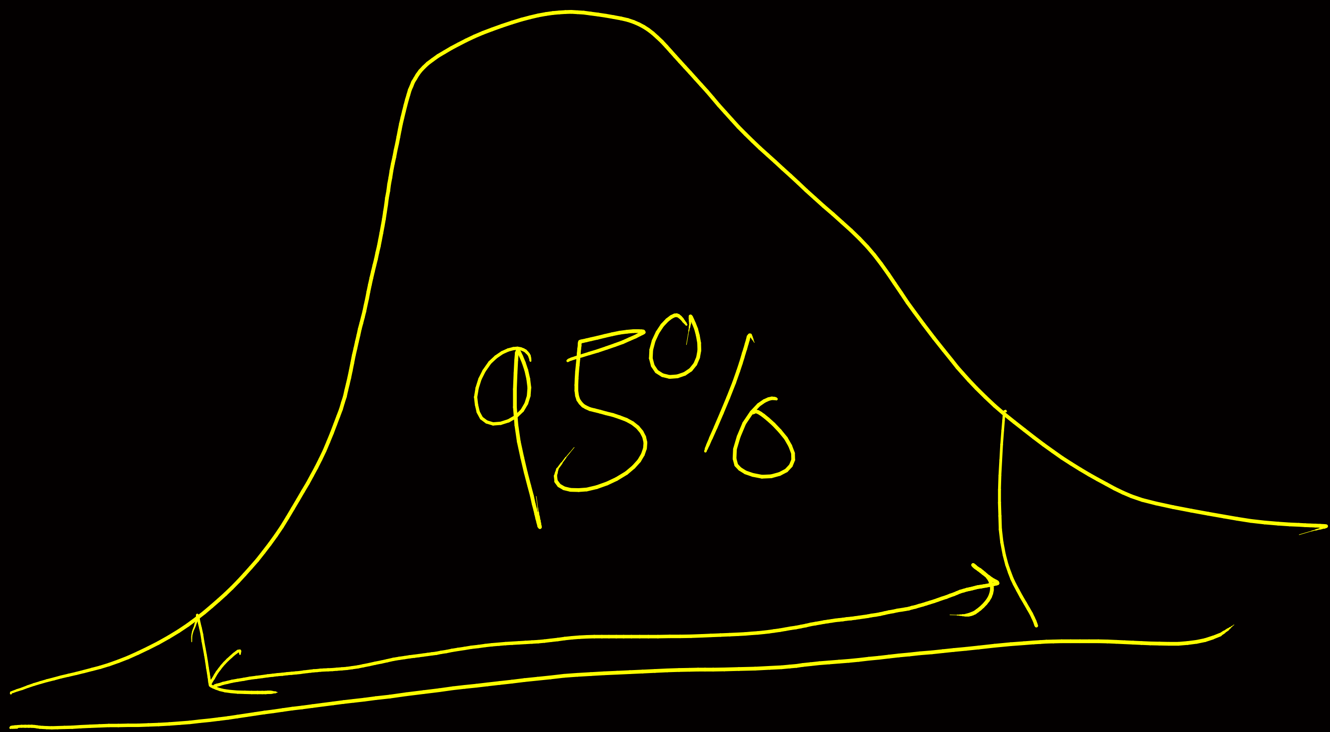
σ

σ^2

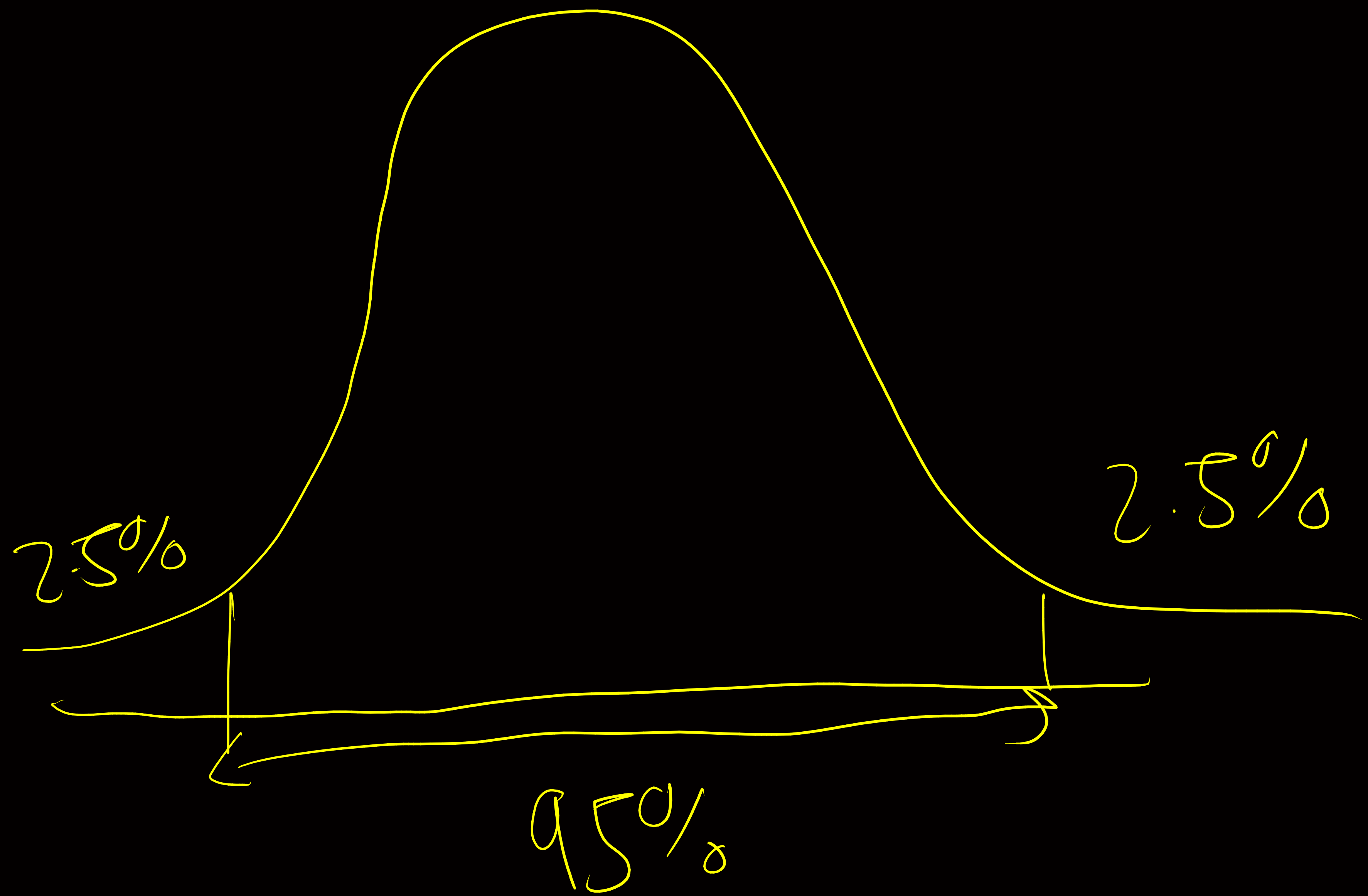
$1/\sigma^2$

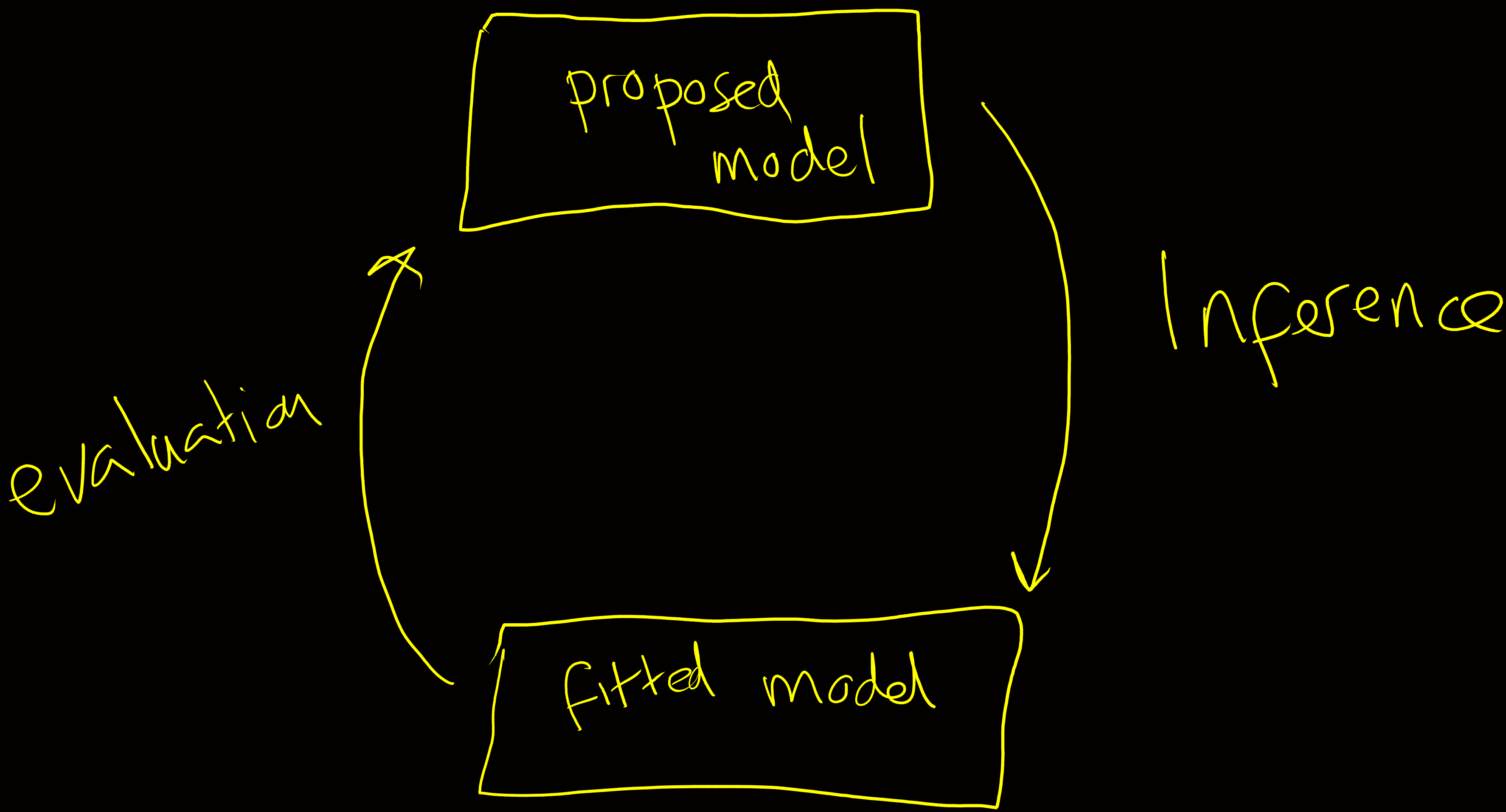
θ

$\log\left(\frac{\theta}{1-\theta}\right)$

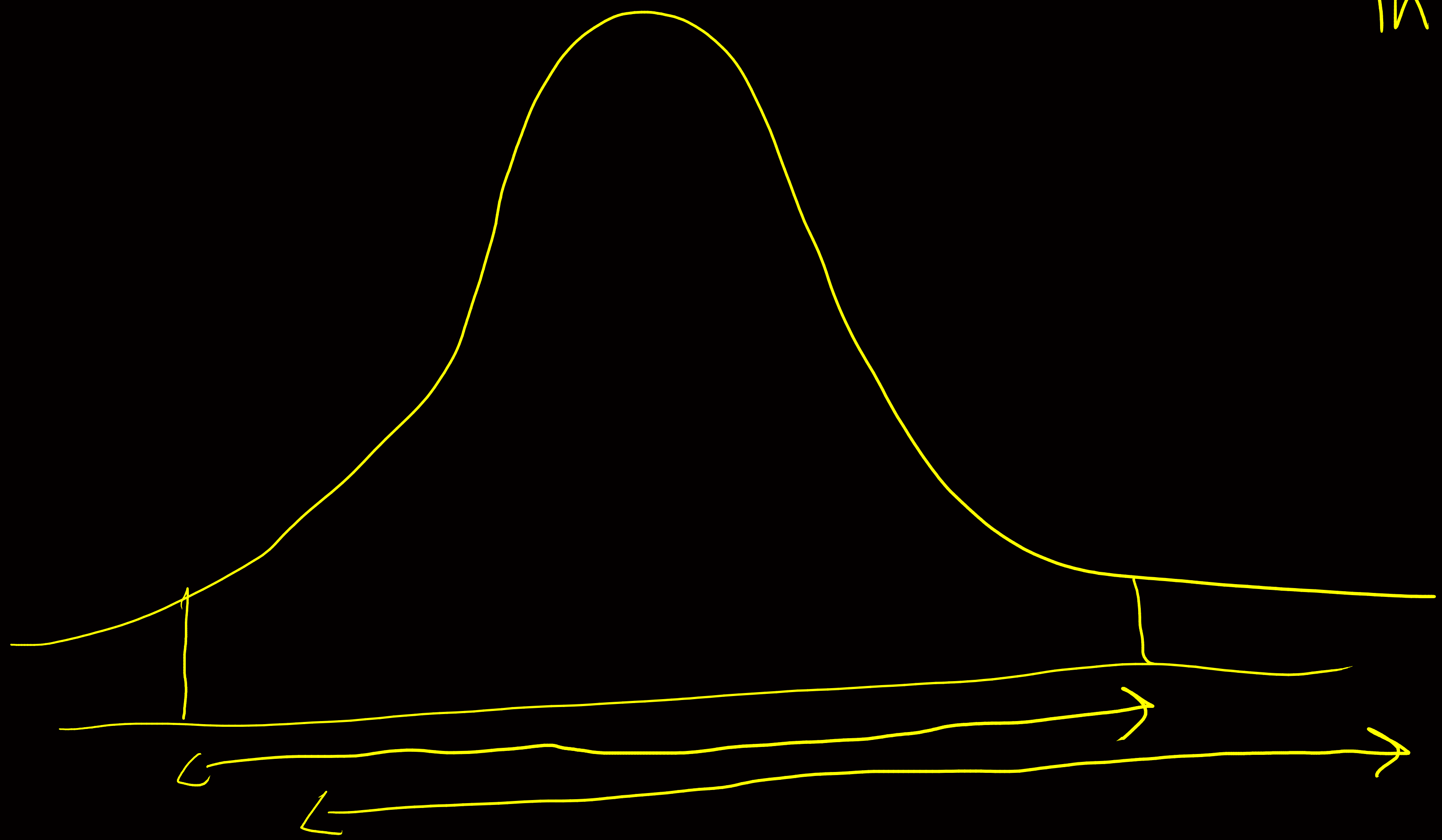


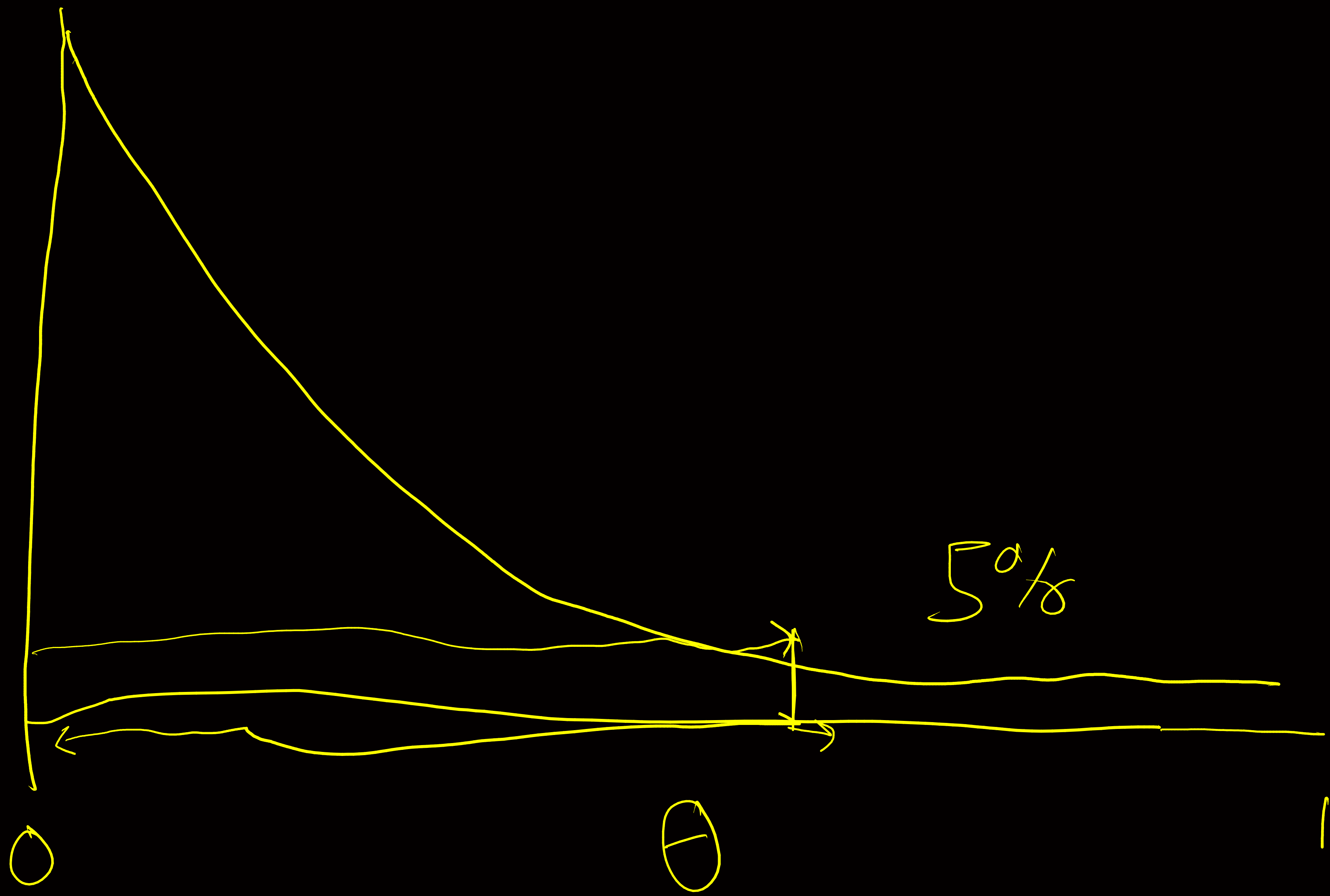
95 HPD


2.5%
2.5%
95%


proposed model
Inference
fitted model
evaluation

95% posterior interval





5%
0
θ
1

$$y_1 \quad y_2 \quad \dots \quad y_n$$

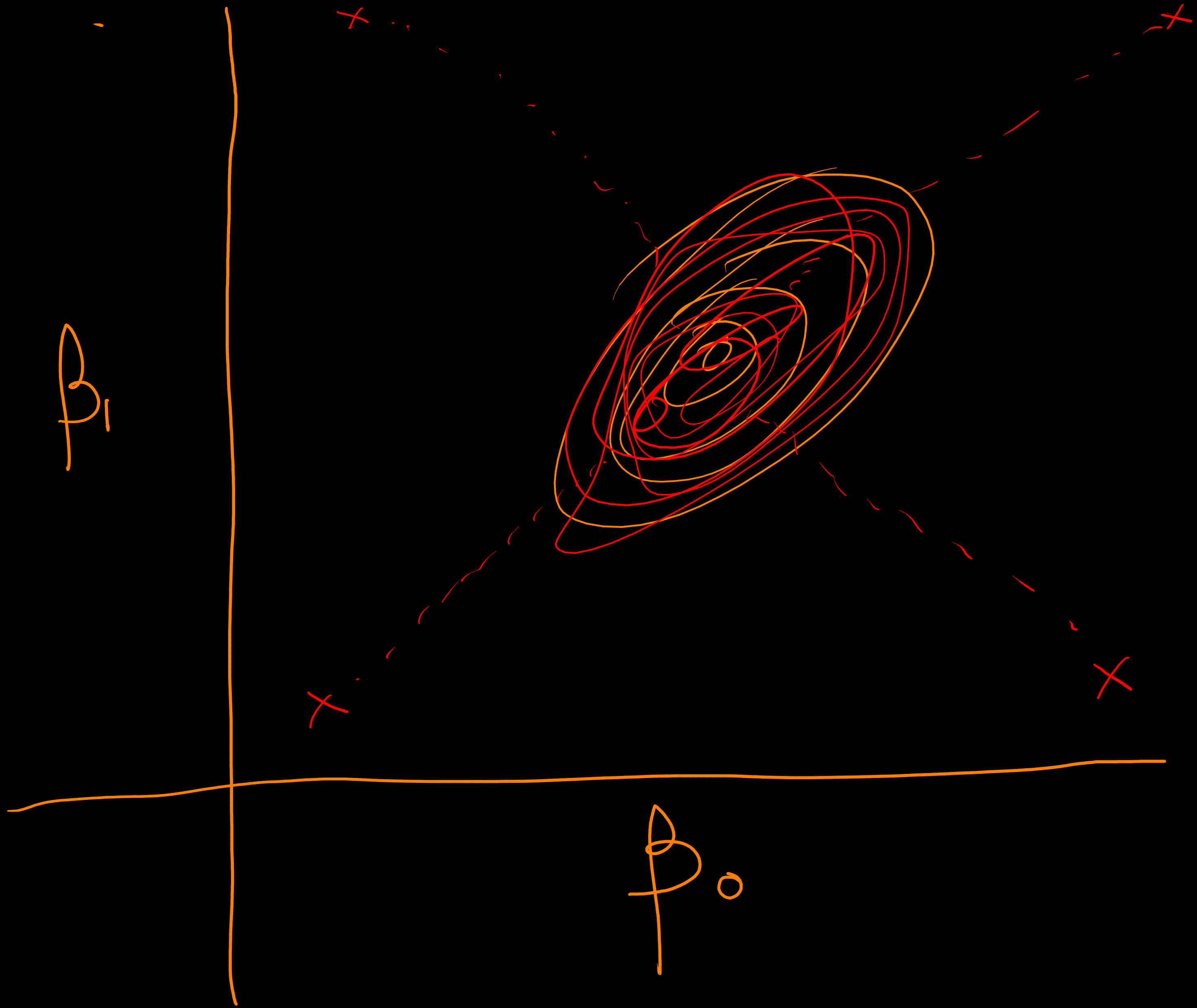
$$x_1 \quad x_2 \quad \dots \quad x_n$$

$$y_i \sim \text{Student}(\mu_i, \sigma)$$

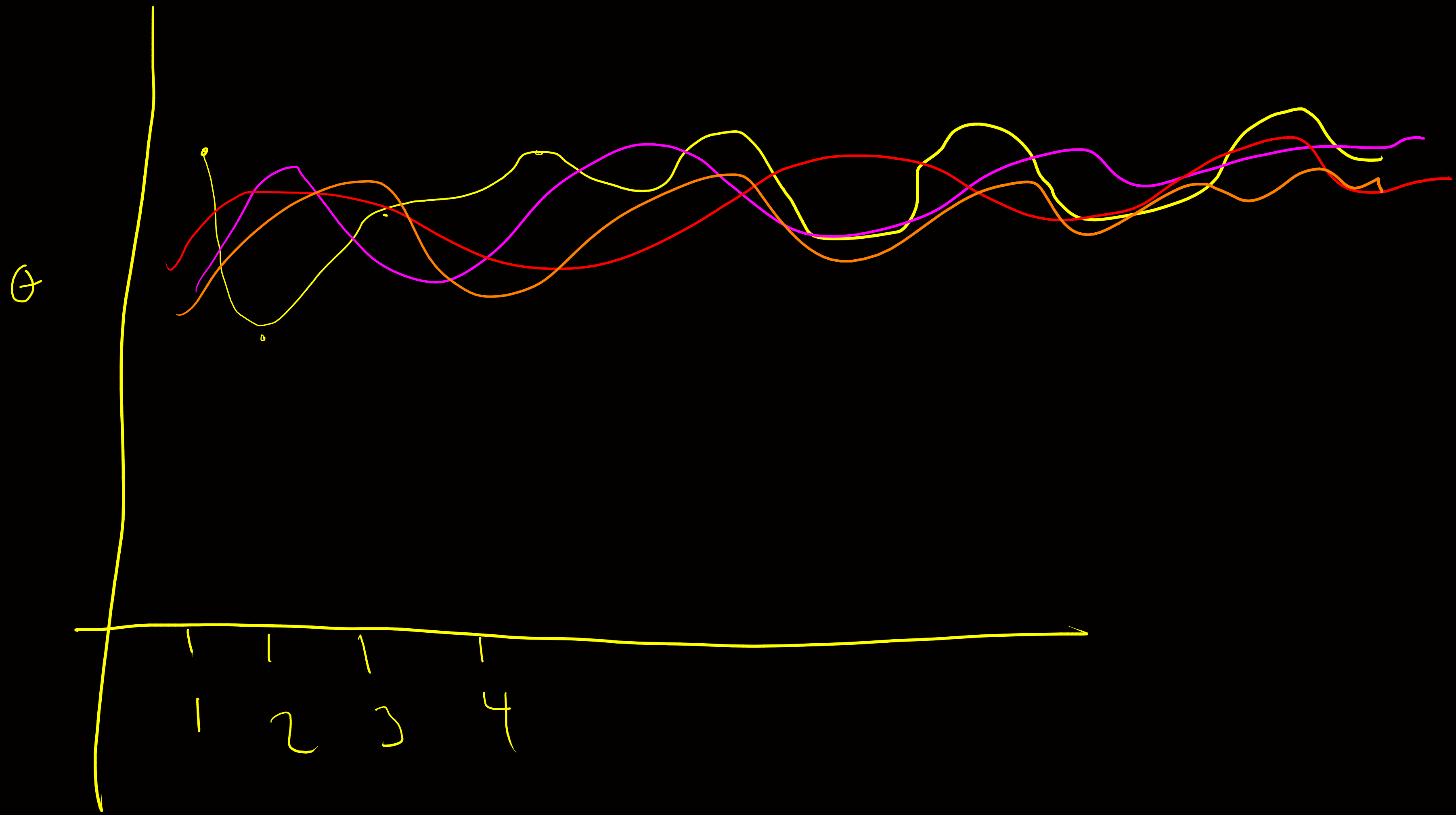
for i in $1 \dots n$

$$y_i \sim N(\mu_i, \underline{\sigma^2})$$

$$\mu_i = \underline{\beta_0} + \underline{\beta_1} x_i$$

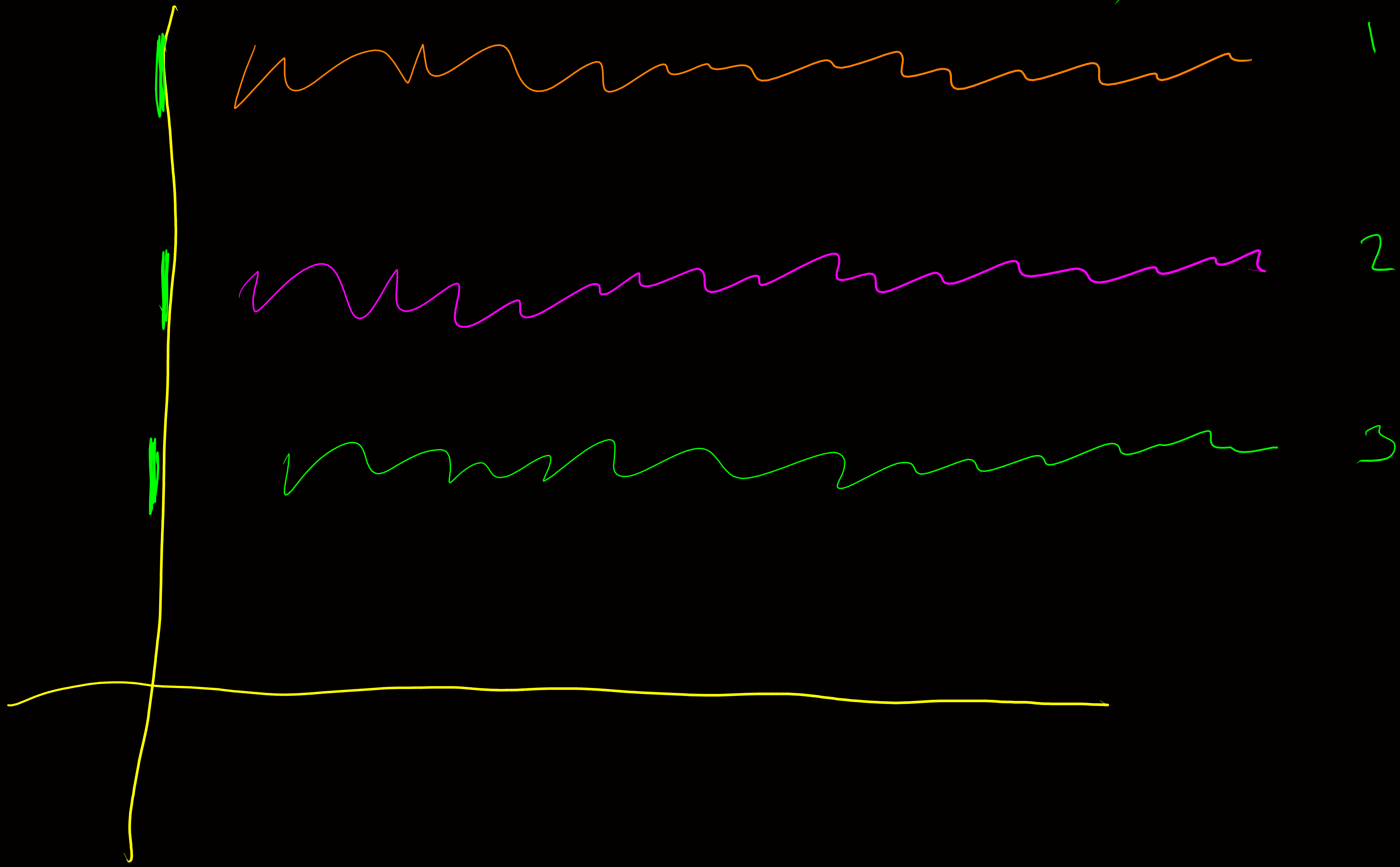


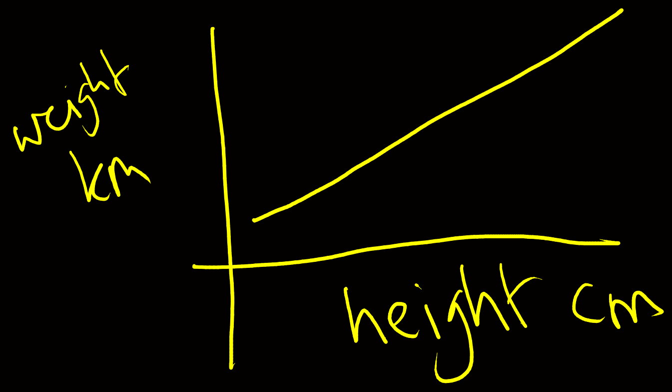
β_1
β_0

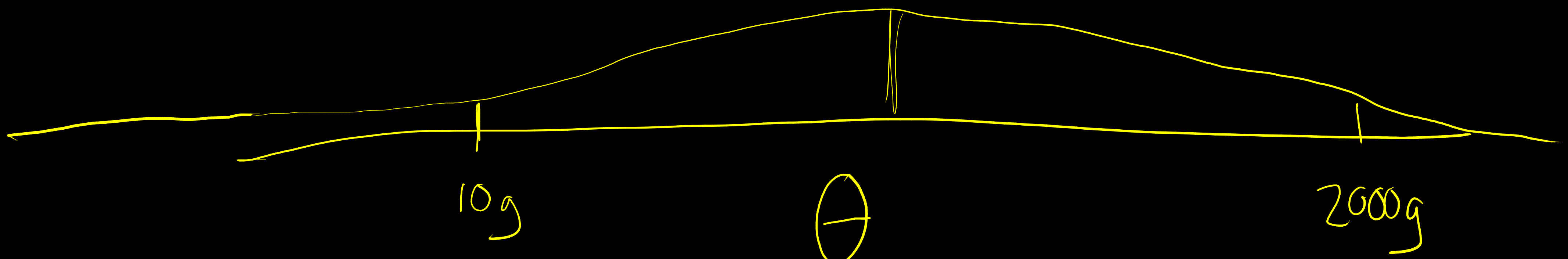


θ
1
2
3
4

Rhat 2, 10




weight
km
height cm

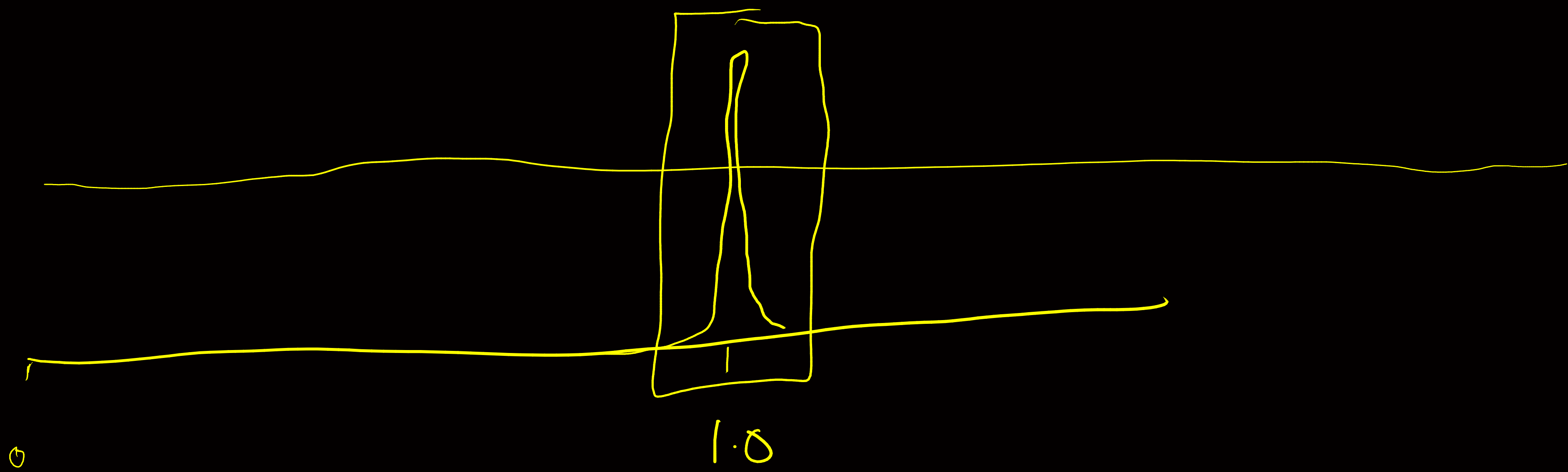

10g
θ
2000g

* change settings
* change priors
* variants of linear models
* model comparison
* generalized linear

  logistic

  Poisson / negbin / zeroinflated

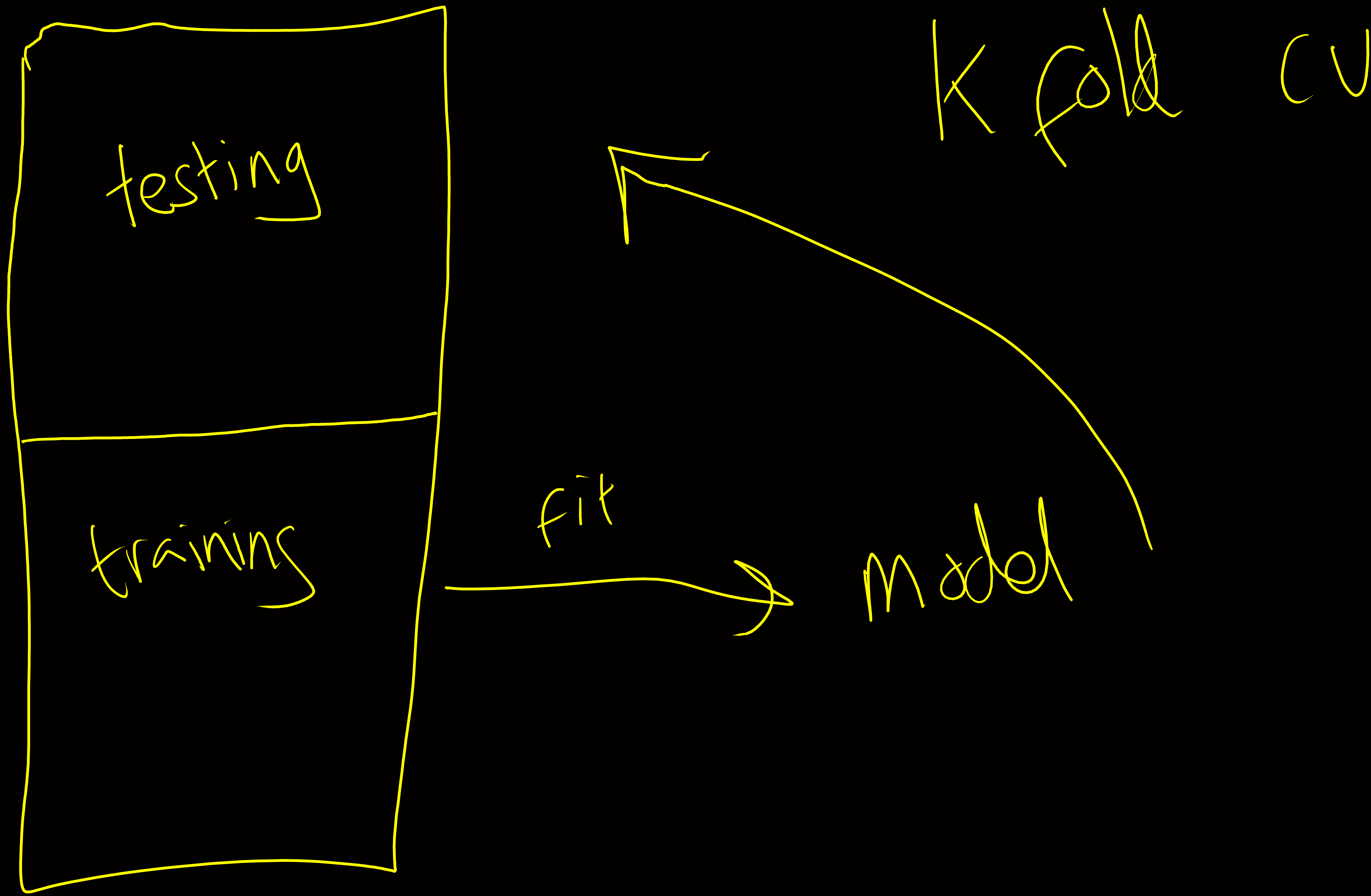
* multilevel model

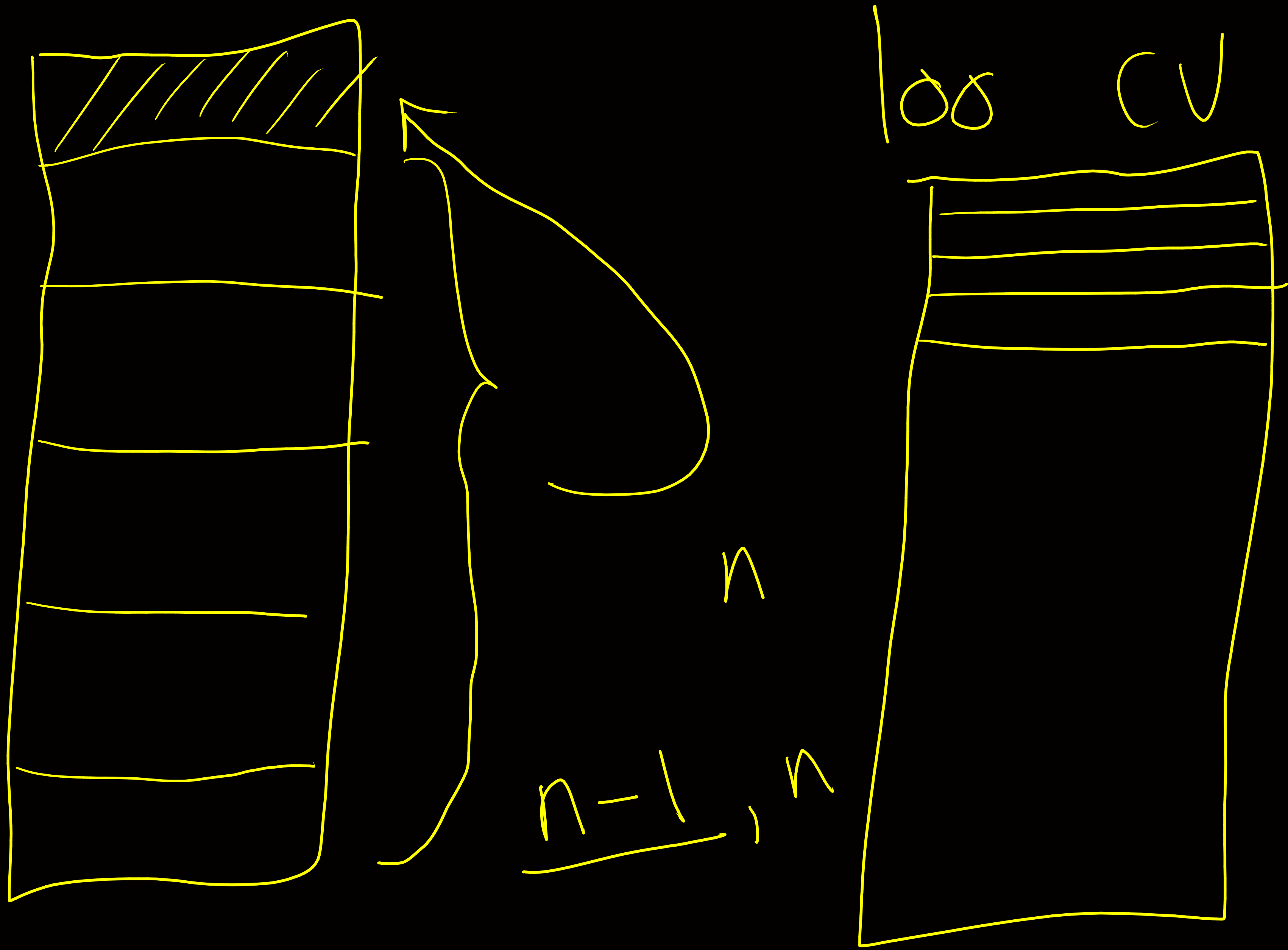

0
1.0

for i in 1, . . . , n

$$y_i \sim t(\nu, \mu_i, \sigma)$$

$$\mu_i = \beta_0 + \sum_{k=1}^{K} \beta_k x_{ki}$$


K fold CV
testing
training
fit
model


100 CV
n
n-1, n

AIC

Akaike Information Criterion

fit data

complexity

bias

variance

$$-2\,LL + 2k$$

$$\approx -2 \log \text{elpd}$$

AICc

$n >> k$

looic

-2 elpd

WAIC

Watanbe Akaike IC

widely applicable

$$M_0 \,, \quad M_1$$

$$D$$

$$p(M_0 \mid D) = \frac{p(D \mid M_0)\, p(M_0)}{p(D \mid M_0)\, p(M_0) + p(D \mid M_1)\, p(M_1)}$$

$$\frac{p(M_0 \mid D)}{p(M_1 \mid D)} = \underbrace{\frac{p(D \mid M_0)}{p(D \mid M_1)}}_{BF} \times \frac{p(M_0)}{p(M_1)}$$

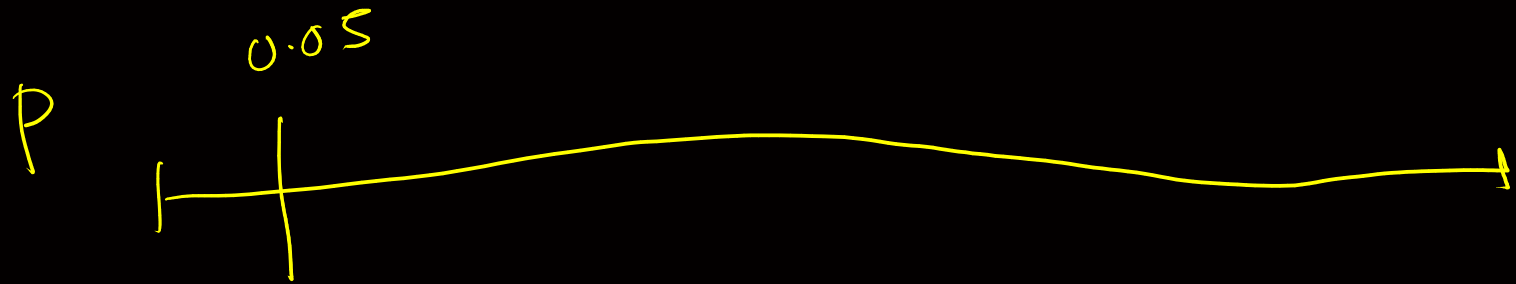
$$p(D \mid M_0) = \underbrace{\int p(D \mid \theta, M_0)\, p(\theta \mid M_0)}_{\text{marginal likelihood}}$$

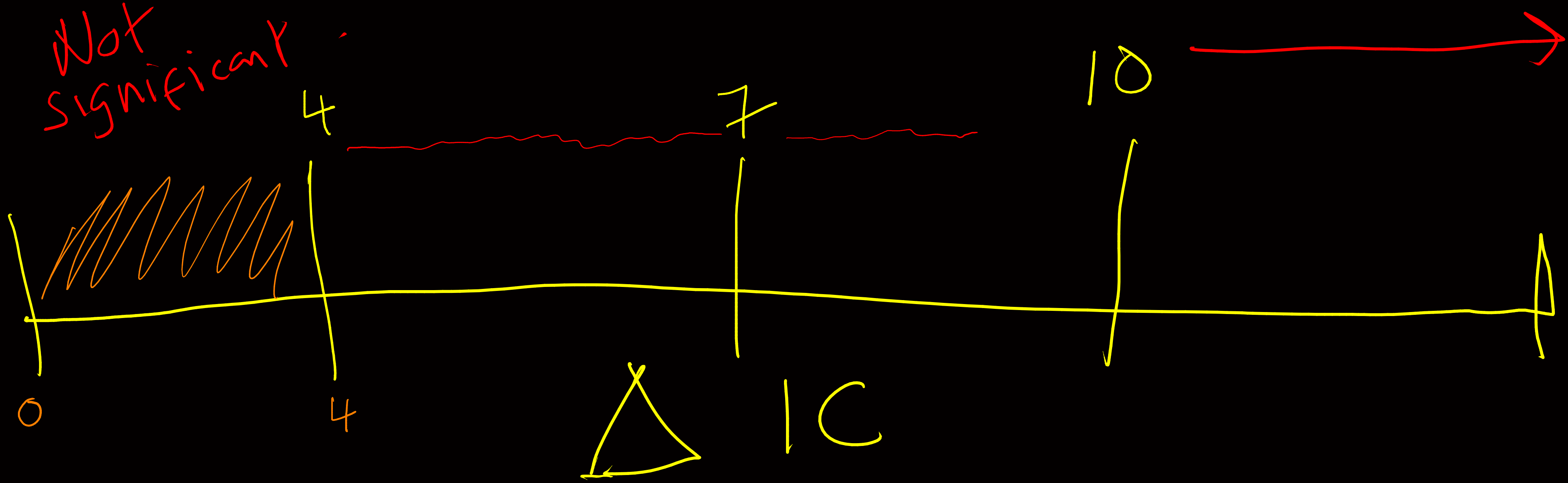
prior predictive density

0.05

0.01

0.001





$$\Delta IC = \left(-2 \log \text{elpd}_1\right) - \left(-2 \log \text{elpd}_i\right)$$

$$= -2 \log \left(\frac{P(\tilde{D} \mid m_1)}{P(D \mid M_2)} \right)$$

$$e^{80}$$

IC

$-2\,LLR$

$$\log\left[\frac{P(D|M_9)}{P(D|M_8)}\right] \simeq 16$$

$\log BF$

$$\frac{P(D|M_9)}{P(D|M_8)} \simeq e^{16}$$

$$\Delta IC = (-2\,LLR_1) - (-2\,LLR_2)$$

$$-2\left(LLR_1 - LLR_2\right)$$

$$-2\log\left(\frac{L_1}{L_2}\right)$$

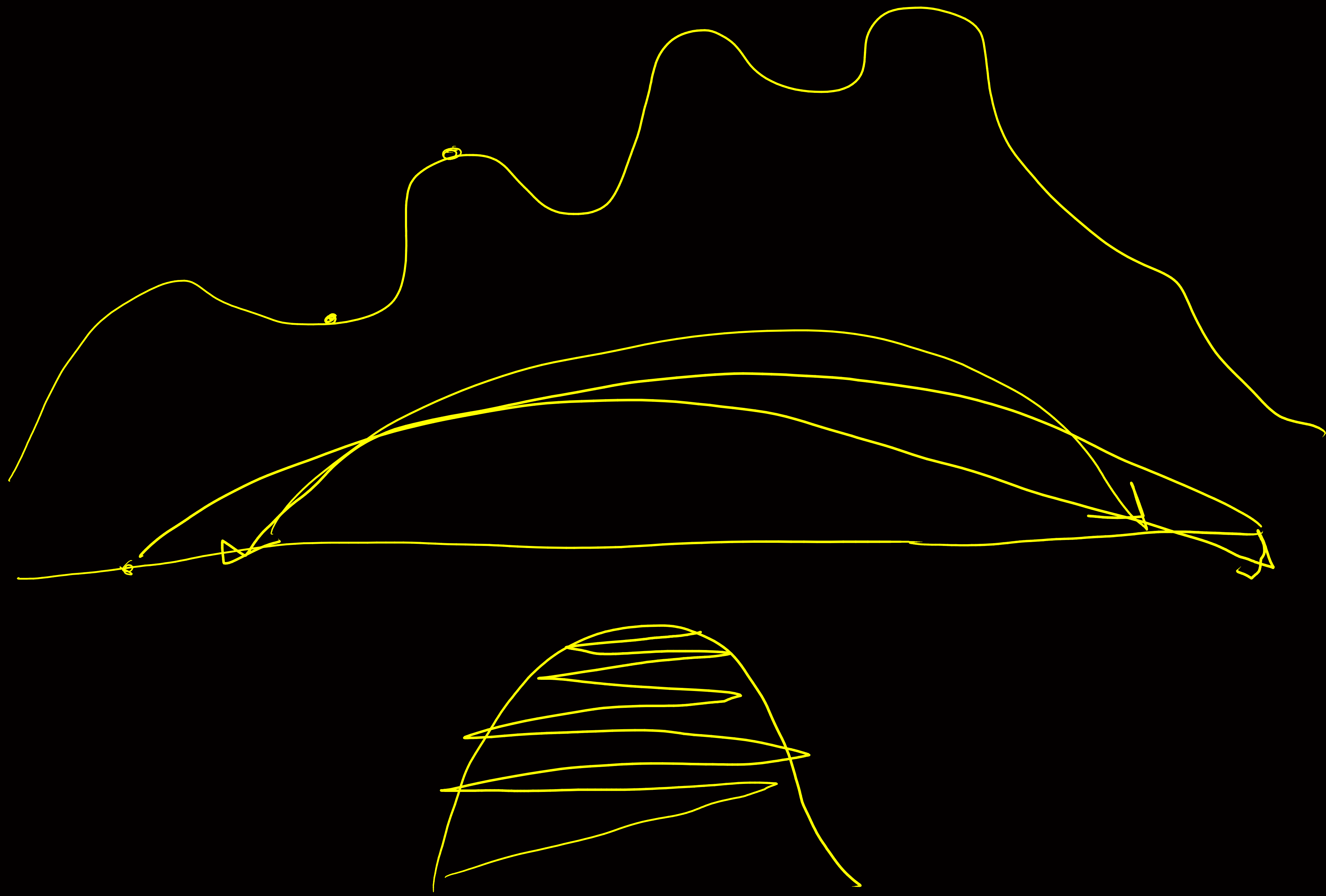
$x_1 \quad x_2 \quad \cdots \quad x_n$

$$\bar{X} = \frac{1}{n}\sum X_i \quad , \quad S_p = \sqrt{\frac{\sum (X_i - \bar{X})^2}{n}}$$

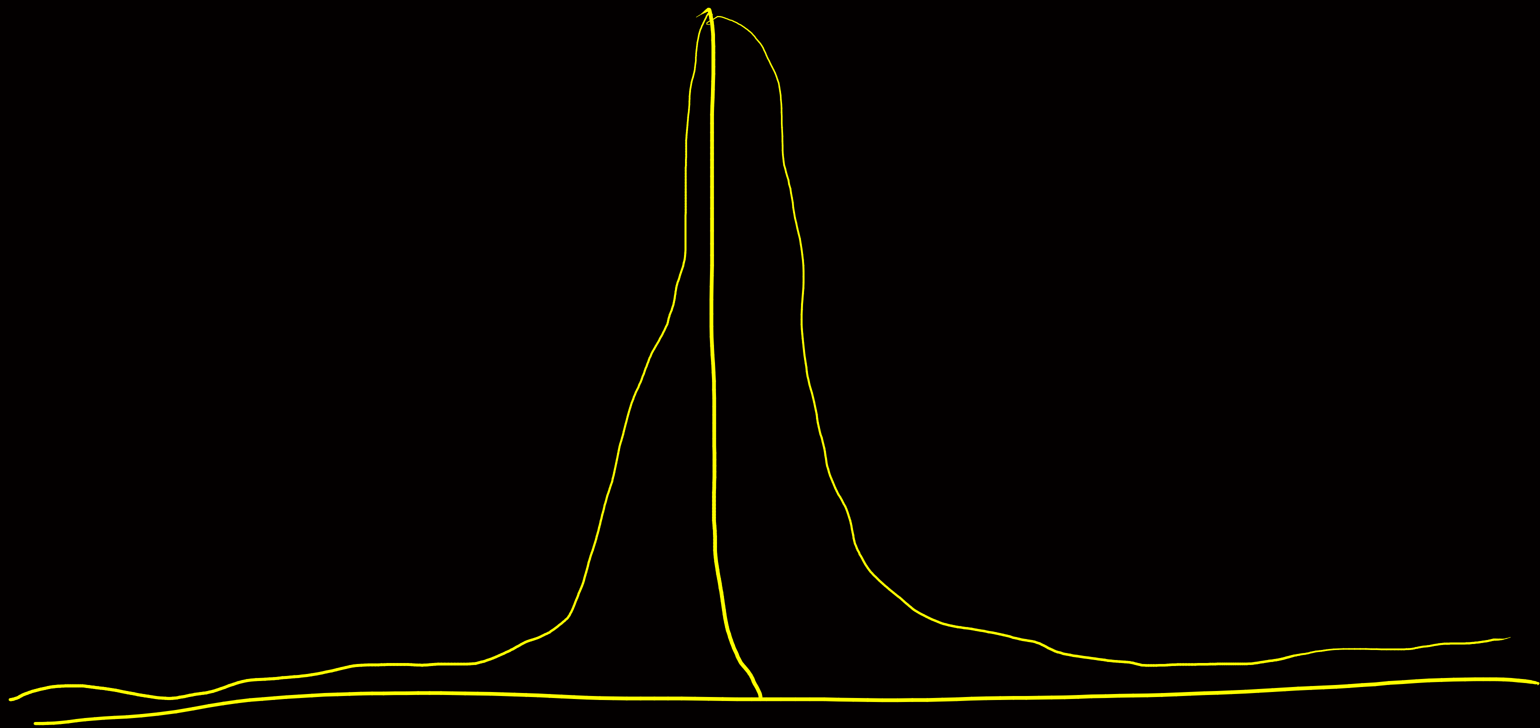
mle for μ mle for σ

$$S = \sqrt{\frac{\sum (x_i - \bar{x})^2}{n-1}}$$

rmle for σ



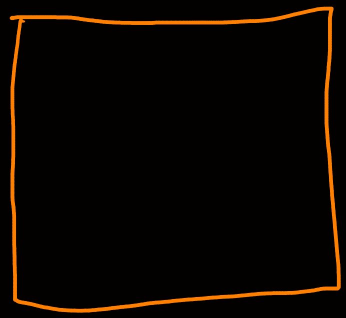
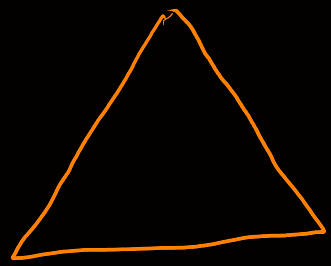
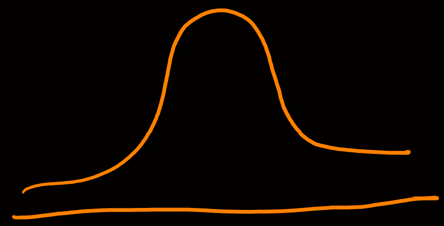
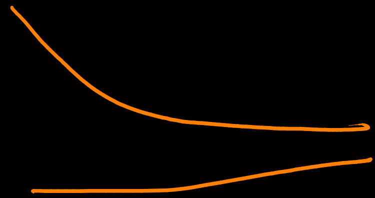

σ
0



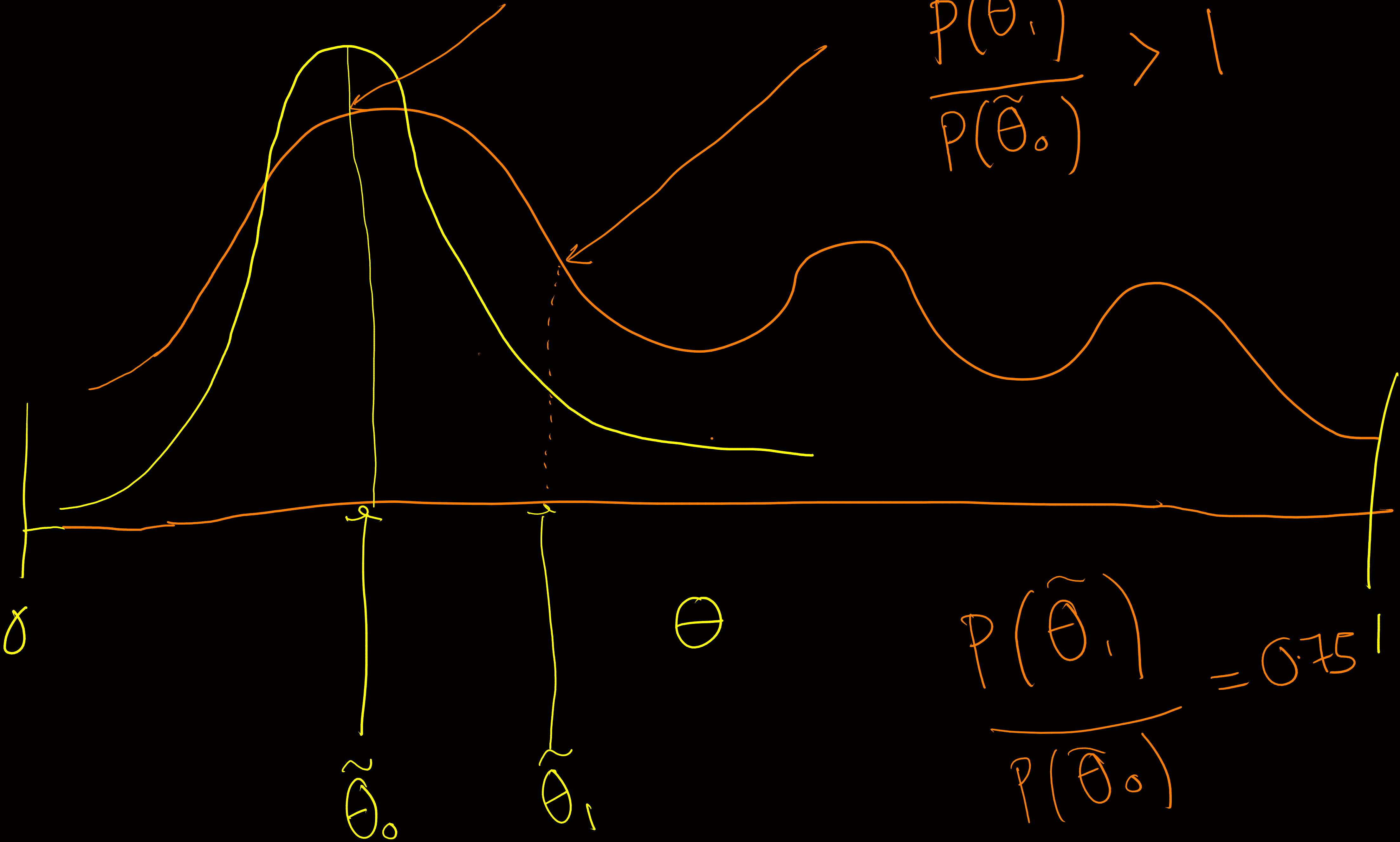
Ca° r

$$\beta_0 + \beta_1 x_1 + \beta_2 x_2$$

brms → Stan → C++ —compiling→ sampler


$$\frac{p(\tilde{\theta}_1)}{p(\tilde{\theta}_0)} > 1$$

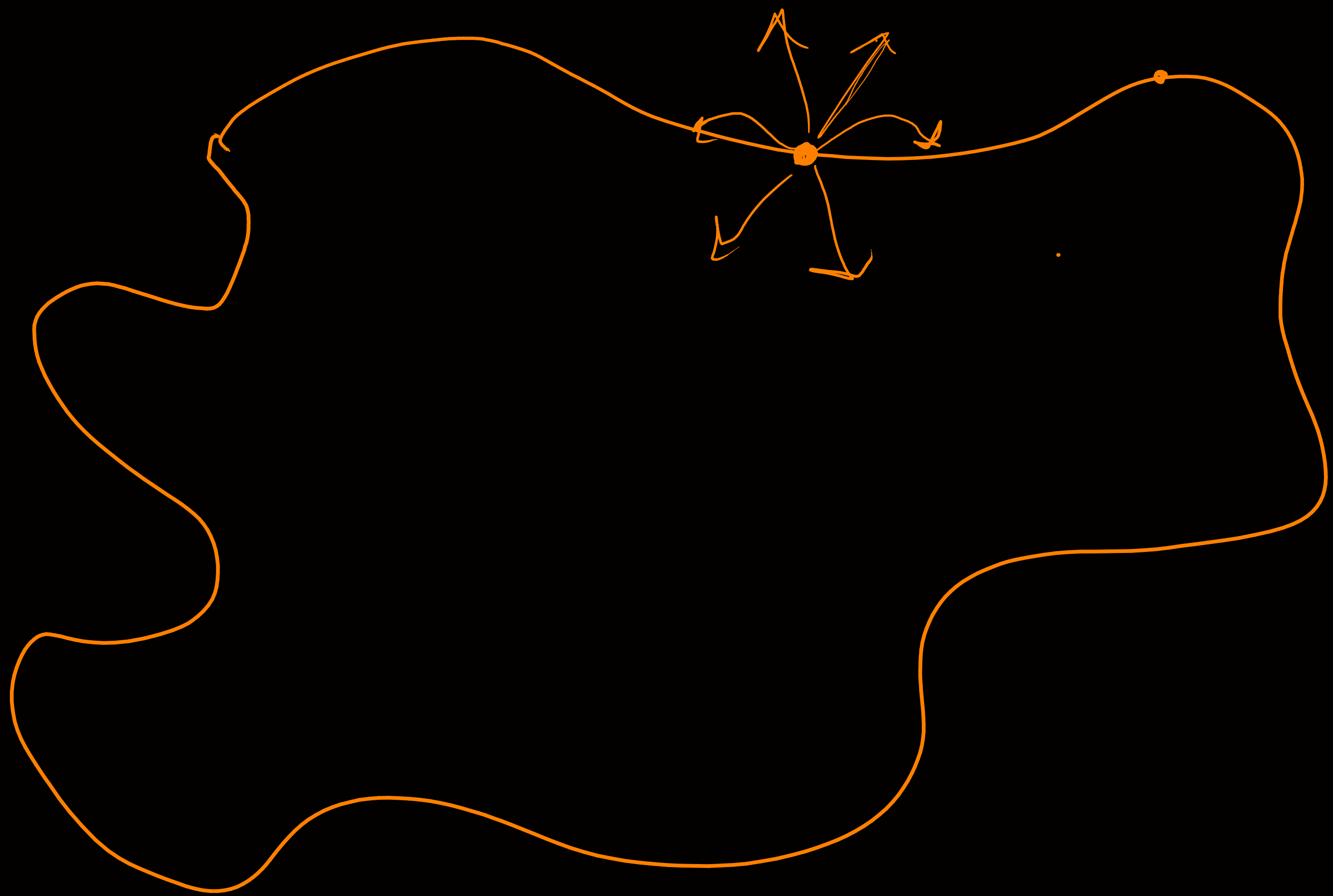


$$\frac{p(\tilde{\theta}_1)}{p(\tilde{\theta}_0)} = 0.75$$

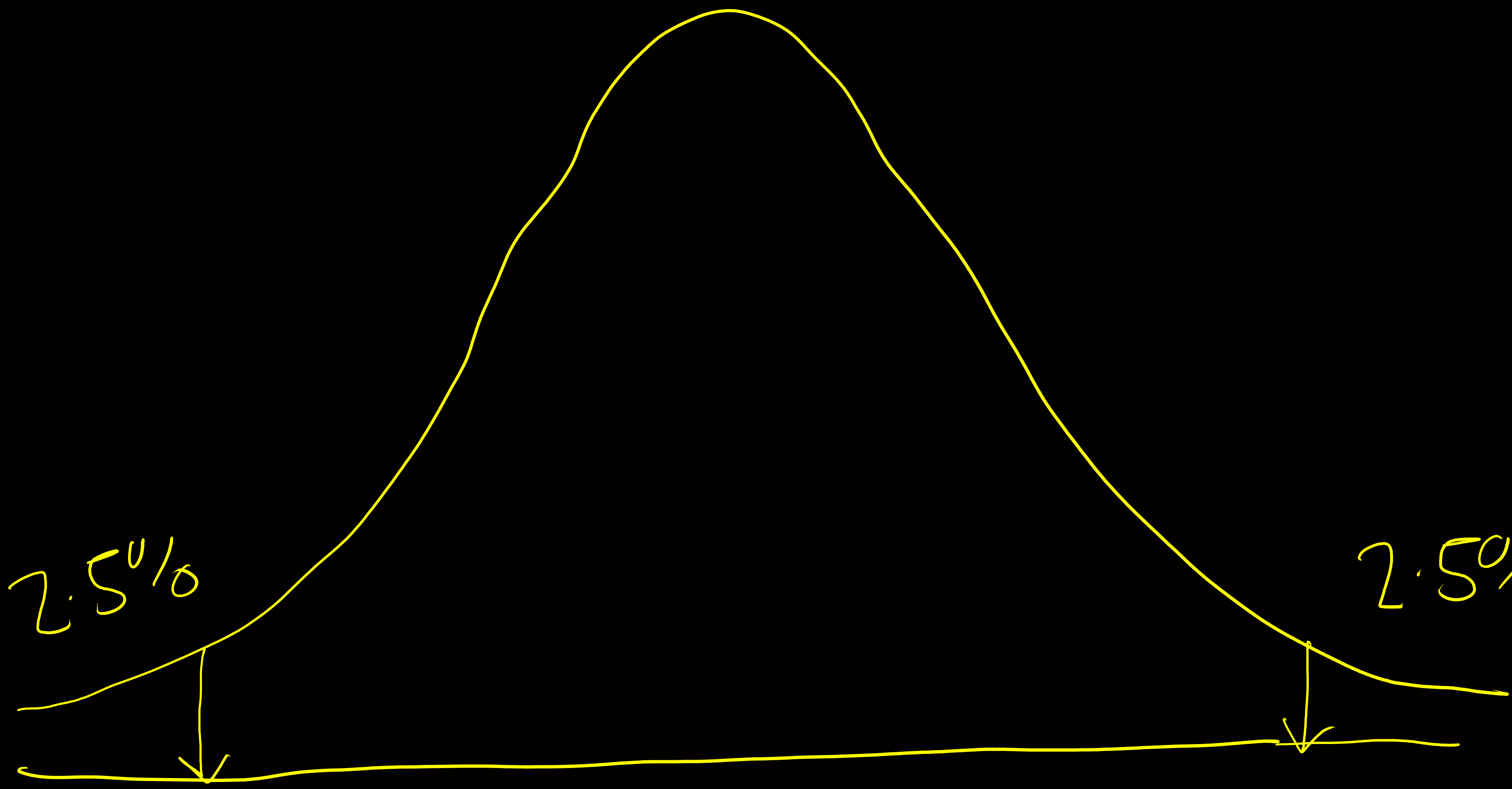
$$\cancel{p(\theta|D)} = \frac{p(D|\theta)\,p(\theta)}{\int p(D|\theta)\,p(\theta)}$$

$\tilde{\theta}_0$ $\tilde{\theta}_1$

$$\frac{p(\tilde{\theta}_1|D)}{p(\tilde{\theta}_1|D)} = \frac{\dfrac{\boxed{p(D|\tilde{\theta}_0)\,p(\tilde{\theta}_0)}}{\cancel{\int p(D|\theta)\,p(\theta)\,d\theta}}}{\dfrac{p(D|\tilde{\theta}_1)\,p(\tilde{\theta}_1)}{\cancel{\int p(D|\theta)\,p(\theta)\,d\theta}}}$$



Typical set



2.5%
2.5%

$$p(x=0) = \int p(x=0 \mid y)\, p(y)\, dy$$