W02 - Probability and inference of parameters

MacKay chapter 3, lecture 10

Sivia chapter 2

"What's Bayesian inference" (a primer SEddy)

Bacterial mutation times

Bacteria can spontaneously mutate

virus sensitive $\longrightarrow$ virus resistant

[Luria + Delbrück, 1943] $\left\{ \begin{array}{l} \to \text{mutations} \checkmark \\ \text{or} \\ \to \text{adaptive immunity} \times \end{array} \right.$

* We want to estimate the (expected) time for a bacterium to mutate

Setup: We observe a bacterial colony for 20 minutes, record times at which we observe bacterium to become resistant

[0, 20] minutes

$N = 6$ became resistant at times

1.2, 2.1, 3.4, 4.1, 7, 11 minutes

assumptions

i) each bacterium mutates independently

ii) mutations do not revert (at least in the time period we are observing

This is an example of a poisson process

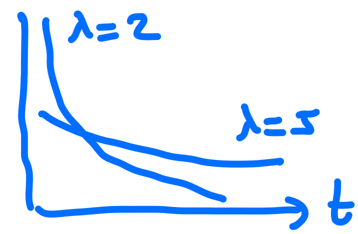
Poisson process

there is a parameter $\lambda > 0$ that controls the rate of events occurring per unit time.

- Time (t) to first event = Exponential dist.

$$P(t) = \frac{1}{\lambda} e^{-t/\lambda}$$

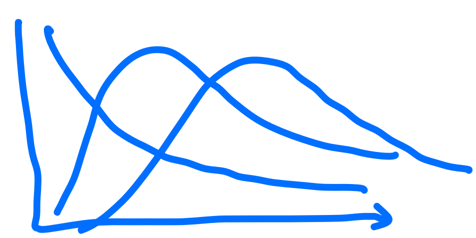
$$\langle t \rangle = \lambda \,,\, \sigma^2 = \lambda^2$$



- Time (t) to r^{th} event = Gamma dist

$$P(t|\lambda, r) = \frac{t^{r-1}}{\lambda^r} \frac{e^{-t/\lambda}}{\Gamma(r)}$$

$$\langle t \rangle = \lambda r \,,\, \sigma^2 = \lambda^2 r$$



- # mutations (n) in time t = Poisson dist

$$P(n|\lambda, t) = \frac{(t/\lambda)^n}{n!} e^{-t/\lambda}$$

$$\langle n \rangle = t/\lambda$$

$$\sigma^2 = t/\lambda$$

Q: Given the data:

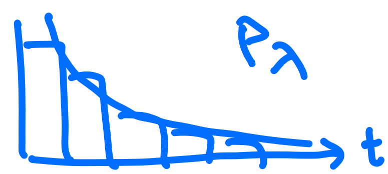
$N=6$

times to mutation $=\{1.2, 2.1, 3.4, 4.1, 7, 11\}$

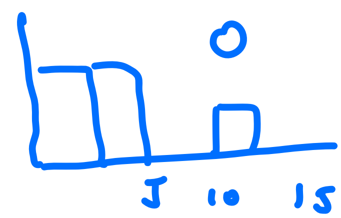
What can we say about the mutation parameter λ?

→ we want to infer λ from the data

how?



① · exponential $P_\lambda = \frac{1}{\lambda} e^{-t/\lambda}$



· data

calculate

$D_{KL}(O \| P_\lambda) \to$ optimize λ

② $\langle t \rangle = \lambda$ take sample mean

$$\lambda \simeq \frac{1}{6}(1.2+2.1+3.4+4.1+7+11)$$

$$= 4.97 \text{ mins}$$

the Bayesian approach

$$\text{data} = \{t_1, t_2, \ldots, t_6\}$$

$$P(\text{data}|\lambda) = P(t_1 t_2 t_3 t_4 t_5 t_6 | \lambda)$$

$$\text{independence} \rightarrow = \prod_{i=1}^{6} P(t_i|\lambda)$$

$$P(t|\lambda) = \begin{cases} \frac{1}{Z(\lambda)} e^{-t/\lambda} & t \in [a,b] \quad (a<b) \\ 0 & \text{otherwise} \end{cases}$$

$$\int_a^b \frac{e^{-t/\lambda}}{Z(\lambda)} dt = 1 \Rightarrow Z(\lambda) = \int_a^b e^{-t/\lambda} dt$$

$$Z(\lambda) = -\lambda e^{-t/\lambda} \Big|_a^b = \lambda \left[e^{-a/\lambda} - e^{-b/\lambda} \right]$$

$$\boxed{P(\text{data}|\lambda) = \frac{e^{-\sum_i t_i/\lambda}}{Z^N(\lambda)}} \qquad N = 6$$

$P(\text{data} \mid \lambda) \longrightarrow P(\lambda \mid \text{data})$

forward probability

- can be measured
- describes the outcome of a random variable

inverse probability

- quantity not directly measurable

Bayes theorem

$$P(AB) = P(BA)$$

$$P(A \mid B) \cdot P(B) = P(B \mid A) \cdot P(A)$$

$$P(B \mid A) = \frac{P(A \mid B) \cdot P(B)}{P(A)}$$

$$\boxed{P(\lambda \mid \text{data}) = \frac{P(\text{data} \mid \lambda) \cdot P(\lambda)}{P(\text{data})}}$$

$$\underbrace{P(\lambda \mid \text{data})}_{\substack{\text{posterior prob} \\ \text{of } \lambda \text{ given data.}}} = \frac{\overbrace{P(\text{data} \mid \lambda)}^{\text{likelihood of } \lambda} \cdot \overbrace{P(\lambda)}^{\text{prior}}}{\underbrace{P(\text{data})}_{\text{evidence}}}$$

$P(\text{data} \mid \lambda) \longrightarrow$ what we know
- depends on data + hypothesis
- "likehood of λ"

or

probability of data given λ

$P(\lambda) \longrightarrow$ prior probability of λ

"not an estimate of λ"

MAX ENT $\rightarrow$ $P(\lambda) = 1$

$P(\text{data}) \longrightarrow$ evidence

does not depend on λ

by marginalization

$$P(\text{data}) = \int_{\lambda'} P(\text{data} \mid \lambda') \cdot P(\lambda')\, d\lambda'$$

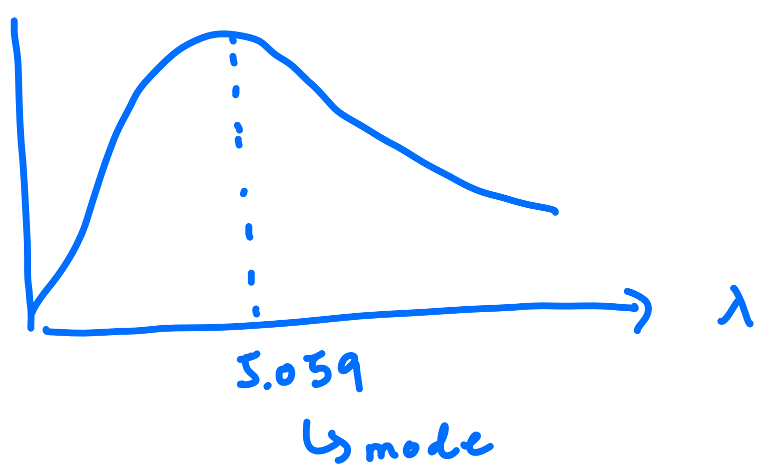
Bacterial mutation times

$$P(\text{data}|\lambda) = \frac{e^{-\sum_i t_i/\lambda}}{Z^6(\lambda)} \qquad \sum_i t_i = 6 \times 4.97$$

$$P(\lambda) = 1 \quad \text{uniform prior.}$$

$$P(\lambda|\text{data}) = \frac{P(\text{data}|\lambda) \cdot P(\lambda)}{P(\text{data})}$$

$$\propto \frac{e^{-\sum_i t_i/\lambda}}{Z^6(\lambda)}$$



Observe the difference between

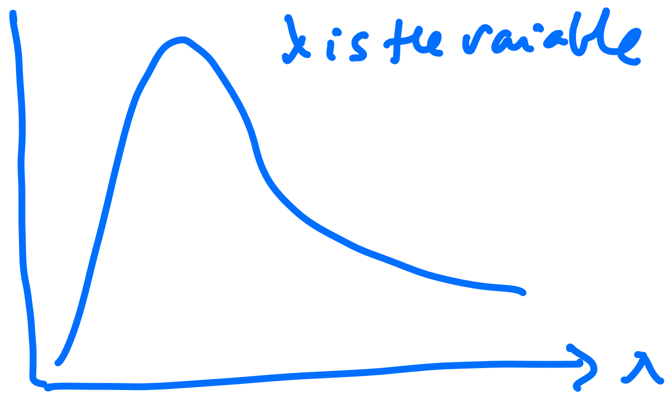
$P(t|\lambda)$

λ is fixed

$$P(\lambda|\text{data}) \propto \prod_{i=1}^{N} P(t_i|\lambda)$$

data are fixed $\{t_i\}$

λ is the variable



As $N \uparrow$ increases, your certainty about the value of λ increases