Problem 3.1

3.1. Using ML estimation to learn parameters

For poisson distribution

$$p(x/\lambda) = \frac{e^{-\lambda}(\lambda)^{x}}{x!}$$

↳ parameter 'λ'.

Estimating 'λ' from MLE :

Likelihood:

$$L(\lambda) = p(\cancel{Y/\lambda}) = \prod_{i=1}^{n} p(y_i/\lambda)$$

$$= \prod_{i=1}^{n} \frac{e^{-\lambda}(\lambda)^{y_i}}{y_i!} = \frac{e^{-N\lambda}(\lambda)^{\Sigma y_i}}{\prod y_i!}$$

Considering log-likelihood does not change anything

$$J(\lambda) = \log(L(\lambda)) = \sum_{i=1}^{n} -\lambda + y_i \log(\lambda) - \log(y_i!)$$

To find the best 'λ' ⇒ need to minimise log-likelihood

$$\Rightarrow \boxed{\frac{\partial J(\lambda)}{\partial \lambda} = 0}$$

$$\frac{\partial J}{\partial \lambda} = \sum_{i=1}^{n} -1 + \frac{x_i}{\lambda} = 0$$

$$\Rightarrow \quad -N + \frac{\sum_{i=1}^{n} x_i}{\lambda} = 0$$

$$\Rightarrow \quad \boxed{\lambda = \frac{\sum x_i}{N}}$$

$\therefore$ 'λ' is the mean of the data, in MLE.

3.2 MAP estimation

⇒ a) Here, I assumed a prior distribution of 'λ' parameter ~~as~~ of each corps as a 'γ'-distribution.

→ Reasons:-

* ~~Let the~~ By seeing the data, we get a rough idea that values at every corps lie in {0, 1, 2, 3, 4} values.

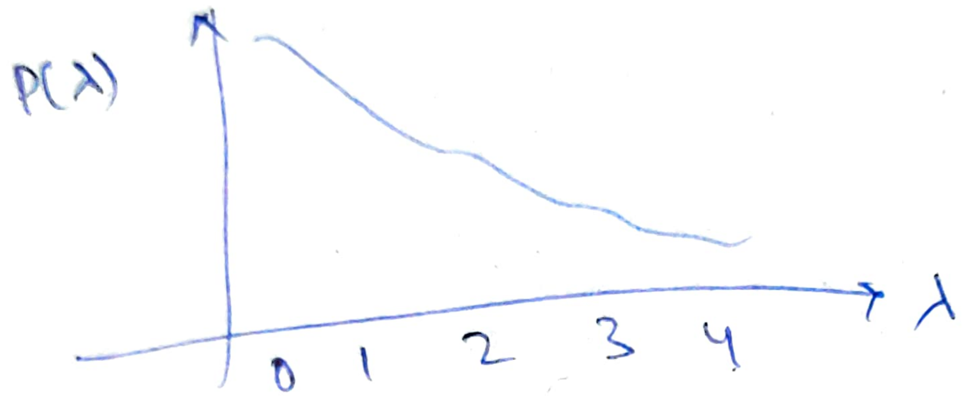
→ Where, we see that ~~corps~~ no. of deaths for each corps is closer to '0'

i.e; $P(0) > P(1) > P(2) > P(3) > P(4)$

↳ Since, we are fitting a poisson distribution, whose mode (high probable value) is floor(λ) ↳ mean.

→ ∴ Predictions are same as 'λ' parameter

Hence, we assume that λ is ~~more~~ such that it will be be equal to {0, 1, 2, 3, 4 ...} with decreasing probability. Something like this



⇒ Mode of 'λ' $\longrightarrow 0$

variance should be more

For γ distribution

$$\text{mode} = \frac{\alpha - 1}{\beta} \rightarrow 0 \qquad \text{var} = \alpha/\beta^2$$

$\therefore \alpha \rightarrow 1$,

$\beta < 1 \rightarrow$ for variance.

(ii) Since counts $\geq 0 \Rightarrow \lambda \geq 0$.

→ For Gamma distribution; this fits exactly.

$$\therefore P(\lambda / \alpha, \beta) = \frac{\beta^{\alpha}}{\Gamma(\alpha)} \lambda^{\alpha-1} e^{-\beta\lambda} \; ; \; \lambda > 0$$

(iii) Considering prior as Gamma, will lead to the aposterior being gamma as well.

$\Rightarrow$ Proof: $P(\lambda | \alpha, \beta) = \dfrac{\beta^{\alpha}}{\Gamma(\alpha)} \lambda^{\alpha-1} \cdot e^{-\beta\lambda} \; ; \; \lambda > 0$

Likelihood: $P(y / \lambda) = \dfrac{e^{-N\lambda} \lambda^{\sum_i x_i}}{\prod x_i!}$

$\therefore P(\lambda / y) \rightarrow$ Posterior.

$P(\lambda / y) \propto P(\lambda) P(y / \lambda) \propto$

$$\propto \left(\frac{\beta^{\alpha}}{\Gamma(\alpha) \prod x_i!} \right) \left(\lambda^{\alpha - 1 + \sum_i^N x_i} e^{-(N+\beta)\lambda} \right)$$

$\sim \Gamma(\tilde{\alpha}, \tilde{\beta})$, where $\tilde{\alpha} = \alpha + \sum x_i$

$\tilde{\beta} = \beta + N$

$\therefore$ Posterior $\sim \Gamma(\tilde{\alpha}, \tilde{\beta})$

$\therefore$ Since the posterior is $\Gamma(\tilde{\alpha}, \tilde{\beta})$

the mode of $\lambda = \dfrac{\tilde{\alpha}-1}{\tilde{\beta}} = \dfrac{\sum x_i + \alpha - 1}{N + \beta}$ ✓

→ Prediction:

→ $Y\_Pred = round(\lambda_{mode}) = \hat{Y}$

RMSE:

Root-mean square error:

$$e = \sqrt{\frac{\sum_{i=1}^{N} (Y_i - \hat{Y}_i)^2}{N}}$$