Ви маєте задачу: передбачити кількість звернень у службу підтримки вашого сайту в певний день. У службу підтримки за день звертається цілочисельна кількість людей, яких зазвичай не більше 7–10, тому ви вирішили використовувати розподіл Пуассона для моделювання таких звернень.

a) Розподіл імовірностей Пуассона має вигляд: $P(y;\lambda) = \dfrac{e^{-\lambda}\lambda^y}{y!}$ ***Покажіть, що розподіл Пуассона належить до експоненціального сімейства і вкажіть, чому дорівнюють*** $b(y), \eta, T(y), a(\eta)$

Відповідь:

Розподіл імовірності належить до експоненціального сімейства, якщо його можна представити у вигляді: $P(y;\eta) = b(y)exp(\eta^\top T(y) - a(\eta))$. Можемо переписати розподіл Пуассона, застосовуючи exp :

$P(y \mid \eta) = exp(log(\dfrac{\lambda^y e^{-\lambda}}{y!})) = exp(ylog\lambda - \lambda - logy!) = \dfrac{1}{y!}exp(ylog\lambda - \lambda)$, де
$\eta = log\lambda, T(y) = y, a(\eta) = exp(\eta), \mu = exp(\eta)$

/---/

b) Якою буде канонічна функція відгуку (canonical response function) для цього розподілу? Ви можете використати той факт, що випадкова величина з розподілом Пуассона з параметром λ має середнє значення λ.

Відповідь:

Оскільки, $\eta = log\lambda \Rightarrow \lambda = e^\eta$

/---/

c) Для навчальної вибірки $\{(x^{(i)}, y^{(i)}); i = 1, \dots m\}$ ***логарифмічна функція правдоподібності (log-likelihood) буде:*** $l(\theta) = logP(y^{(i)} \mid x^{(i)};\theta)$ ***Виведіть похідну*** $\frac{\partial}{\partial\theta_j} l(\theta)$ ***та сформулюйте правило оновлення ваги*** θ_j ***методом стохастичного градієнтного підйому, якщо y має розподіл Пуассона та канонічну функцію відгуку.***

Відповідь:

Розподіл можна записати у вигляді $P(y^{(i)} \mid x^{(i)}, \theta) = \frac{1}{y!}exp(y^{(i)}\theta^\top x^{(i)} - e^{\theta^\top x^{(i)}})$, тому можемо виразити ймовірність у вигляді $P(y^{(n)} \mid (x^{(n)}, \theta) = \displaystyle\prod_{i=1}^{n} \dfrac{exp(y^{(i)}\theta^\top x^{(i)} - e^{\theta^\top x^{(i)}})}{y^{(i)}!}$

Обчислимо likelihood функцію $l(y \mid x, \theta) = \sum_{i=1}^{n}(y^{(i)}\theta^{\top}x^{(i)} - exp(\theta^{\top}x^{(i)}) - log(y^{(i)}!))$

Візьмемо часткову похідну, щоб максимізувати likelihood функцію

$$\frac{\partial l(y|x,\theta)}{\partial\theta} = \sum_{i=1}^{n}(y^{(i)}x^{(i)} - x^{(i)}exp(\theta^{\top}x^{(i)})) = \sum_{i=1}^{n}x^{(i)}(y^{(i)} - exp(\theta^{\top}x^{(i)})) = \sum_{i=1}^{n}x^{(i)}(y^{(i)} - \widehat{y^{(i)}})$$

$$\theta = \theta - \alpha\nabla_{\theta}l(\theta) \Rightarrow \theta_j = \theta_j + \alpha(x^{(i)}(y^{(i)} - exp(\theta^{\top}x^{(i)}))$$

```
In [ ]:
```