1.A. MAP MLE Common distributions

Friday, December 27, 2019 4:13 PM

mle: $\text{argmax}_{(\theta)}\ P(D|\theta)$

map: $\text{argmax}(\theta)\ P(\theta|D) = \boxed{\dfrac{P(D|\theta)P(\theta)}{\cancel{P(D)}}}$

Find mle (maximum likelihood estimator) for the following params

1). Probability of success p in Bernoulli model

example dataset: 1, 1, 0, 1, 1 ; 1, 0, 1, 1, 0

Let X represent the Bernoulli R.V. X_i is i.i.d.

$X = X_1, X_2, X_3 \dots X_n$.

Prob mass function of a Bernoulli is $\begin{cases} p & x=1 \\ (1-p) & x=0 \end{cases}$

equivalently, $p^x(1-p)^{x-1}$ where $x = 0, 1$

$f(x_i|p) = p^x(1-p)^{1-x_i}$

1) $l(X_1, X_2, X_3, \dots X_n|p) = \prod_{i=1}^{n} f(x_i|p)$ (independence)

2) $\sum_{i=1}^{n} \ln[f(x_i|p)]$ (Log transf)

3) $\sum_{i=1}^{n} \ln[p^{x_i}(1-p)^{1-x_i}]$ (PMF of Bernoulli)

4) $= \sum_{i=1}^{n} x_i \ln p + \sum_{i=1}^{n}(1-x_i)\ln(1-p)$ (rules of Log)

5) $\text{argmax}\ l(X_1, X_2, \dots X_n|p) =$

$\text{argmax}\left[\sum_{i=1}^{n} x_i \ln p + \sum_{i=1}^{n}(1-x_i)\ln(1-p)\right]$

$\frac{d}{dp}\left[\sum_{i=1}^{n} x_i \ln p + \sum_{i=1}^{n}(1-x_i)\ln(1-p)\right] = 0$

6) $\frac{\sum_{i=1}^{n} x_i}{p} + (-1)\frac{\sum_{i=1}^{n}(1-x_i)}{1-p} = 0$

7) $\frac{\sum_{i=1}^{n} x_i}{p} = \frac{\sum_{i=1}^{n}(1-x_i)}{1-p}$

8) $(1-p)\sum_{i=1}^{n} x_i = p\sum_{i=1}^{n}(1-x_i)$

$= \sum_{i=1}^{n} x_i - p\sum_{i=1}^{n} x_i = p\sum_{i=1}^{n}(1-x_i)$

$\sum_{i=1}^{n} x_i = p\left(\sum_{i=1}^{n} x_i + \sum_{i=1}^{n}(1-x_i)\right)$

$\sum_{i=1}^{n} x_i = p\left(\sum_{i=1}^{n} 1\right)$

$\sum_{i=1}^{n} x_i = p(n)$

$$\boxed{\frac{1}{n}\sum_{i=1}^{n} x_i = \hat{p}}$$

=> $\hat{p}_{mle}$ is the sample mean of the Bernoulli distribution

2. Probability of success p in Binomial(n, p)

The pmf for a binomial distribution is $P(x) = \binom{n}{x}p^x(1-p)^{n-x}$ where $x = 0 \dots n$

Let X be a Binomial R.V. w/ params n and p. Let $X_1, \dots X_m$ be iid from X.

$l(X_1 \dots X_m|n, p) = \prod_{i=1}^{m}\binom{n}{x}p^x(1-p)^{n-x}$ (independence)

$= \cancel{\sum_{i=1}^{m}\ln\binom{n}{x_i}} + \frac{\sum_{i=1}^{m} x_i}{p} - \frac{\sum_{i=1}^{m} n-x_i}{1-p}$

maximizing over p

$= \frac{\sum_{i=1}^{m} x_i}{p} = \frac{\sum_{i=1}^{m} n - \sum_{i=1}^{m} x_i}{1-p}$

$= (1-p)\sum_{i=1}^{m} x_i = p\sum_{i=1}^{m} n - p\sum_{i=1}^{m} x_i$

$\hat{p}_{mle} = \frac{\sum_{i=1}^{m} x_i}{mn}$

3. Probability of success p in a Geometric model

Let X be i.i.d. from a Geometric r.v. w/ parameter p.

The pmf for a Geometric R.V is $f(x) = p(1-p)^{x-1}$.

$l(X_1 \dots X_n|p) = \prod_{i=1}^{n} p(1-p)^{x_i-1}$ (independence)

Log-likelihood:

$\sum_{i=1}^{n} \ln p + (x_i - 1)\ln(1-p)$

$\text{argmax}_p = \frac{\partial}{\partial p}\left[\sum_{i=1}^{n} \ln p + (x_i-1)\ln(1-p)\right] = 0$

$= \frac{\sum_{i=1}^{n} 1}{p} = \frac{\sum_{i=1}^{n} x_i - 1}{1-p}$

$= (1-p)\sum_{i=1}^{n} 1 = p\sum_{i=1}^{n} x_i - 1$

$n - \cancel{np} = \sum_{i=1}^{n} p x_i - \cancel{np}$

$\frac{n}{\sum_{i=1}^{n} x_i}$

4. Intensity λ in Poisson(λ) models

Let X be a poisson R.V w/ $X_1 \dots X_n$ i.i.d from X. The pdf of a poisson is $f(x) = \frac{e^{-\lambda}\lambda^x}{x!}$

$l(X_1 \dots X_n|\lambda) = \prod_{i=1}^{n}\frac{e^{-\lambda}\lambda^{x_i}}{x_i!}$ (independence)

Log likelihood:

$\sum_{i=1}^{n} -\lambda + x_i \ln(\lambda) \ln\left[\prod_{i=1}^{n}\frac{1}{x_i!}\right]$ $\quad c = \ln\left[\prod_{i=1}^{n}\frac{1}{x_i!}\right]$

$\sum_{i=1}^{n} -\lambda + \sum_{i=1}^{n} x_i \ln(\lambda). \ln(c)$

$\frac{\partial}{\partial\lambda}\left[\sum_{i=1}^{n} -\lambda + \sum_{i=1}^{n} x_i \ln(\lambda) \ln(c)\right]$

$= -n + \frac{\sum_{i=1}^{n}\lambda_i}{\lambda} + 0 = 0$

$n = \frac{\sum_{i=1}^{n} x_i}{\lambda}$ => $\boxed{\hat{\lambda}_{mle} = \frac{\sum_{i=1}^{n} x_i}{n}}$