1. $q \rightarrow$ the number of boot failure before Staring

Priya · S · Kulkarni
1002088875

$\rightarrow$ follows geometric distribution.

$$P_q(k) = (1-q)^{k-1} q \quad \text{- hypothesis space.}$$

1a) performance function:

1. parameter - Maximum likelihood estimation
2. Most likely parameter - maximum apostierior probability

MLE

$\rightarrow$ Analytic optimization of gradient ascent

$$P(D/q) = \prod_{i=1}^{n} P_q(k_i)$$

$\rightarrow$ optimization criterion.

$$q^* = \operatorname{argmax}_q P(D/q)$$

Computing the derivative.

$$\frac{\partial P(D/q)}{\partial q} = \frac{\partial}{\partial q}\left(\prod_i^{n} P_q(k_i)\right) = \frac{\partial}{\partial q}\left(\prod_i^{n} P(k_i/q)\right)$$

$\Rightarrow$ ~~puting~~ log on both side:

$$\frac{\partial}{\partial q}\left(\ln \prod_i^{n} P(k_i/q)\right) = \frac{\partial}{\partial q} \sum \ln P(k_i/q)$$

$$= \sum \frac{\partial}{\partial q} \ln P(k_i/q)$$

$$= \sum \frac{\partial}{\partial q} \ln\left[\cancel{P_q}(1-q)^{k-1} * q\right]$$

$$= \sum \frac{\partial}{\partial q}\left[\ln(1-q)^{k-1} + \ln q\right]$$

$$= \sum \frac{\partial}{\partial q}\left[(k-1)\ln(1-q) + \ln q\right]$$

$$= \sum\left[\frac{\partial}{\partial q}(k-1)\ln(1-q) + \frac{\partial}{\partial q}\ln q\right]$$

$$= \sum \left[(k-1) \frac{1}{1-q} (-1) + \frac{1}{q} \right] = \sum \left[\frac{1-k}{1-q} + \frac{1}{q} \right]$$

$$= \sum \left[\frac{1}{1-q} - \frac{k}{1-q} + \frac{1}{q} \right] = \frac{n}{1-q} - \frac{\sum k_i}{1-q} + \frac{n}{q}$$

therefore: $\frac{\partial}{\partial q} \ln P(D/q) = \frac{n}{1-q} - \frac{\sum k_i}{1-q} + \frac{n}{q}$

Equating to 0

$$\frac{n}{1-q} - \frac{\sum k_i}{1-q} + \frac{n}{q} = 0$$

$$\frac{-\sum k_i}{1-q} = -\frac{n}{q} - \frac{n}{1-q}$$

$$\frac{\sum k_i}{1-q} = \frac{n}{q} + \frac{n}{1-q}$$

$$\frac{\sum k_i}{1-q} = \frac{n(1-q) + nq}{q(1-q)}$$

$$\frac{\sum k_i}{\cancel{(1-q)}} = \frac{n - \cancel{nq} + \cancel{nq}}{q\cancel{(1-q)}}$$

$$\sum k_i = \frac{n}{q}$$

$$\boxed{q = \frac{n}{\sum k_i}}$$ optimial Solution — (1)

(1b) Data: 1, 3, 1, 2, 1, 1, 2, 2, 3, 3

$n = 10$

$$q = \frac{n}{\sum k_i} = \frac{n}{(1+3+1+2+1+1+2+2+3+3)}$$

$$q = \frac{10}{19} = \underline{\underline{0.52631}}$$

(1c) Beta distribution

$$P_{\alpha,\beta}(q) = \frac{q^{\alpha-1}(1-q)^{\beta-1}}{B(\alpha,\beta)}$$

$$\Rightarrow \operatorname{argmax}_q P(q/D) = \operatorname{argmax}_q \frac{P(D/q)\,P(q)}{P(D)}$$

ignore $P(D)$ (does not depend on q)

$$q_{MAP} = \operatorname{argmax}_q P(D/q) * P(q)$$

$$= \operatorname{argmax}_q \left[\prod_i^n P(K_i/q)\right] * P(q)]$$

Apply log function.

$$\ln(q_{MAP}) = \operatorname{argmax}_q \log P(q/D)$$

$$= \operatorname{argmax}_q \sum_{i \in K}^{n} [\underbrace{\log P(K_i/q)}_{\text{geometric distribution}} + \underbrace{\log P(q)}_{\text{Beta distribution}}]$$

$\Rightarrow$ differentiating & equating to 0

From equation (1) in 1a

$$\frac{\partial}{\partial q} P(q|D) = \sum_{i=1}^{n} \left[\underbrace{\frac{\partial}{\partial q}\left(\log \prod_i^n P(K_i/q)\right)}_{(1)} + \frac{\partial}{\partial q}(\log P(q))\right]$$

$$= \sum_{i=1}^{n} \left[\left[\frac{n}{1-q} - \frac{\Sigma K_i}{1-q} + \frac{n}{q}\right] + \frac{\partial}{\partial q}\log P(q)\right] \quad (2)$$

Consider $\frac{\partial}{\partial q}\log P(q) = \frac{\partial}{\partial q}\log\left[\frac{q^{\alpha-1}(1-q)^{\beta-1}}{B(\alpha,\beta)}\right]$

$$= \frac{\partial}{\partial q}\left[\log q^{\alpha-1} + \log(1-q)^{\beta-1} - \log B(\alpha,\beta)\right]$$

$$= \frac{\partial}{\partial q}\left[(\alpha-1)\log q + (\beta-1)\log(1-q) - \log B(\alpha,\beta)\right]$$

$$= \frac{\partial}{\partial q}\left[(\alpha-1)\log q + (\beta-1)\log(-q+1) - \underbrace{\log B(\alpha,\beta)}\right]$$

$$= \frac{\partial}{\partial q}(\alpha-1)\log q + \frac{\partial}{\partial q}(\beta-1)\log(-q+1) - \underbrace{\frac{\partial}{\partial q}\log B(\alpha,\beta)}_{=0}$$

$$= \frac{(\alpha-1)}{q} + \frac{\beta-1}{-q+1}(-1) - 0$$

$$= \frac{\alpha-1}{q} + \frac{(\beta-1)(-1)}{(1-q)} = \frac{\alpha-1}{q} + \frac{(1-\beta)}{(1-q)}$$

$$= \frac{\alpha-1}{q} + \frac{1}{1-q} - \frac{\beta}{1-q} = \frac{\alpha}{q} - \frac{1}{q} + \frac{1}{1-q} - \frac{\beta}{1-q} \quad \text{--- (3)}$$

Substituting (3) in (2)

$$\frac{\partial}{\partial q}P(q/D) = \left[\frac{n}{1-q} - \frac{\sum k_i}{1-q} + \frac{n}{q}\right] + \sum\left[\frac{\alpha}{q} - \frac{1}{q} + \frac{1}{1-q} - \frac{\beta}{1-q}\right]$$

$$= \frac{n}{1-q} - \frac{\sum k_i}{1-q} + \cancel{\frac{n}{q}} + \frac{n\alpha}{q} - \cancel{\frac{n}{q}} + \frac{n}{1-q} - \frac{n\beta}{1-q}$$

$$= \frac{2n}{1-q} - \frac{\sum k_i}{1-q} + \frac{n\alpha}{q} - \frac{n\beta}{1-q}$$

Equate to 0.

$$\frac{2n}{1-q} - \frac{\sum k_i}{1-q} + \frac{n\alpha}{q} - \frac{n\beta}{1-q} = 0$$

$$\frac{2n - \sum k_i - n\beta}{1-q} + \frac{n\alpha}{q} = 0.$$

$$\frac{\sum k_i + n\beta - 2n}{1-q} = \frac{n\alpha}{q}$$

$$\frac{\sum k_i + n\beta - 2n}{n\alpha} = \frac{1-q}{q}$$

$$\frac{\sum k_i + n\beta - 2n}{n\alpha} = \frac{1}{q} - 1$$

$$\frac{\sum k_i + n\beta - 2n}{n\alpha} + 1 = \frac{1}{q}$$

$$\frac{\sum k_i + n\beta - 2n + n\alpha}{n\alpha} = \frac{1}{q}$$

$$q = \frac{n\alpha}{\sum k_i + n\beta - 2n + n\alpha}$$

$$\boxed{q = \frac{n\alpha}{\sum k_i - 2n + n(\beta+\alpha)}}$$

Substituting the data values.

$\Sigma k_i = 19 \quad n = 10 \quad \alpha = 4.04 \quad \beta = 4.04$

$$q = \frac{10 \times 4.04}{19 - 2\times10 + 10(4.04 + 4.04)}$$

$$= \frac{40.4}{19 - 20 + 80.8} = \frac{40.4}{79.8} = \underline{\underline{0.50626}}$$

(2a) KNN algorithm is the estimation/prediction of class based on k-nearest neighbors.

<u>1st step</u>: Calculate cartesian distance for each datapoint of test data against training set.

X = Training set of 12 rows.

Y = Test set for class attributes prediction.

Cartesian distance formula $= \sum_{\substack{i \\ i \in c}}^{n} \sqrt{(X_i - Y_i)^2}$

where c is the features of training set.

n=4. there are 4 features

$$= \sqrt{(X_{height} - Y_{height})^2 + (X_{diameter} - Y_{diameter})^2 + (X_{weight} - Y_{weight})^2 + (X_{hue} - Y_{hue})^2}$$

<u>test set has 4 datapoints</u>

P1(0.1008532587158 8, 0.1034766537008 7, 0.6600005512705 4, 3.1061777 201591)

P2(0.0975208056293 66, 0.1201052695695, 0.75, 1.4520706957674)

P3(0.0769730867619 57, 0.0886224896283 88, 0.1060494742654 9, 3.504459418 7921)

P4(0.1184351404545 485, 0.15, 0.3883971487186 3, 5.6111022334215 7)

* <u>P1 & 1st datapoint from training data.</u>

D1 = (0.1178200605301 43, 0.1383652867038 1, 0.3768383115946 2, 3.2559432114899)

$$= \sqrt{(0.1178200053014 3 - 0.100853258715 88)^2 + (0.138365286703 81 - 0.103476653700 87)^2 + (0.376838311594 12 - 0.660000551270 54)^2 + (3.255943211489 9 - 3.106177720159 1)^2}$$

$= 0.322669557867 8$

Similarly all distances are computed.

2nd step: The K-nearest neighbour are classified with K minimum distance & choosing class with highest occurrences.

The following table shows cartesian distance between point 1, point 2, point 3 and point 4 with 12 training data points

Training data	point 1	point 2	point 3	point 4
Plastic	0.32266956	1.84226795	0.37383938	2.35550154..
Plastic	2.7938514..	3.44839608	1.43046965	0.73408212
Metal	0.33525726	1.8959992	0.34889146	2.29794636
Metal	1.3082227	2.91122947	0.78835544	1.34650569
Plastic	0.55159997	1.25181483	0.90431374	2.96854671
Ceramic	0.56691994	1.09834815	1.15525449	3.08929374
Metal	2.11322768	0.7367102	2.44393625	4.55616494
Ceramic	0.3328421	1.37585618	0.80813182	2.81003962
Ceramic	0.87342832	0.78413413	1.40653167	3.39666661
Plastic	1.09294919	2.74229686	0.76169145	1.43830546
Ceramic	1.46750984	0.26648641	1.91900233	3.97656176
Plastic	1.43127784	0.73818276	1.71439447	3.82950435
K=1	Plastic	Ceramic	Metal	Plastic
K=3	ambiguous.	ambiguous	Plastic ~~ambiguous~~	Plastic
K=5	ambigious	Ceramic	ambigous	Plastic ~~ambiguous~~