We have $x = (x_1, x_2, x_3 \cdots x_k) \sim \text{Multi}(p_1, p_2 \cdots p_k)$

$$\Pr\left(x = (x_1, x_2 \cdots x_k) \middle| p\right) = \frac{n!}{x_1! x_2! x_3! \cdots x_k!} \prod_{i=1}^{k} p_i^{x_i}$$

$$x_i \in (0, \cdots n)$$

$$\sum_{i=1}^{k} x_i = n \quad \text{and} \quad \sum_{i=1}^{k} p_i = 1$$

Taking log likelihood (Doesn't alter monotonocity)

$$\log(\Pr(x = (x_1, x_2 \cdots x_k) | p)) = L(x|p)$$

$$L(x|p) = \log n! - \sum_{i=1}^{k} \log(x_i!) + x_i \sum_{i=1}^{k} \log(p_i).$$

For MLE (using Langrange Multiplier)

$$\frac{\delta L(x|p_i)}{\delta p_i} = \gamma \frac{\delta \sum_{i=1}^{k} p_i}{\delta p_i}$$

$$\frac{x_i}{p_i} = \gamma$$

$$p_i = \frac{x_i}{\gamma}$$

Scanned with CamScanner

Using constraint

$$\sum_{i=1}^{K} p_i = \sum_{i=1}^{K} \frac{x_i}{\gamma}$$

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

$$\boxed{\hat{p}_i = \frac{x_i}{n}}$$

$$\text{MLE} \Rightarrow \hat{p}_i = \frac{x_i}{n}$$

Scanned with CamScanner

1. Bayes menthod of Estimation

We have $x = (x_1, x_2, x_3 \ldots\ldots x_k) \sim \text{multi}(p_1, p_2, p_3 \ldots p_k)$

$$Pr(x|p) = n! \prod_{i=1}^{k} \frac{p_i^{x_i}}{x_i!}$$

By Bayes theorem we can write

$$f(p|x) = \frac{Pr(x|p)\, f(p)}{\int Pr(x|p)\, f(p)\, dp} \longrightarrow (i)$$

$Pr(x|p) \Rightarrow$ Data function

$f(p|x) \Rightarrow$ Posterior Distribution

$f(p) \Rightarrow$ Prior distribution

$\int Pr(x|p) f(p)\, dp \Rightarrow$ uniquelly defined marginal distribution.

For multinomial data function we have a Drichlet distribution as conjugate prior this also implies Posterior distribution as Drichlets distribution.

Hence we have drichlet prior $f(p)$

$$p \sim \mathrm{Dir}(\alpha_1, \alpha_2 \ldots \alpha_k) \quad \alpha_i > 0 \text{ for } i = 1, 2 \ldots k$$

$$f(p) = \frac{\Gamma\left(\sum_{i=1}^{k} \alpha_i\right)}{\prod_{i=1}^{k} \Gamma \alpha_i} \prod_{i=1}^{k} p_i^{\alpha_i - 1}$$

where $p_i \in (0,1)$ $\sum_{i=1}^{k} p_i = 1$

By (i) we have,

$$f(p|x) \propto \Pr(x|p) \times f(p)$$

$$f(p|x) \propto \prod_{i=1}^{k} p_i^{x_i} \prod_{i=1}^{k} p_i^{\alpha_i - 1}$$

$$\boxed{f(p|x) \propto \prod_{i=1}^{k} p_i^{(x_i + \alpha_i - 1)}}$$

Hence $f(p|x)$ is a drichlet distribution with parameter $(x_i + \alpha_i)$

Hence posterior distribution ~~of~~ p

$$f(p|x) = \frac{\Gamma\left(\sum_{i=1}^{k} x_i + \alpha_i\right)}{\prod_{i=1}^{k} \Gamma(x_i + \alpha_i)} \prod_{i=1}^{k} p_i^{x_i + \alpha_i - 1}$$

$p_i \in (0,1)$ and $\sum_{i=1}^{k} p_i = 1$

Sol → 2nd

We have mean of Posterior distribution

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

writing posterior mean as convex combination of MLE and prior mean.

$$\text{prior mean} = \frac{\alpha_i}{\sum_{i=1}^{n} \alpha_i}$$

$$\text{MLE} = \frac{x_i}{n}$$

$$\frac{\alpha_i + x_i}{\sum_{i=1}^{k} \alpha_i + n} = \lambda \frac{\alpha_i}{\sum_{i=1}^{k} \alpha_i} + (1-\lambda)\left(\frac{x_i}{n}\right)$$

Solving we get $\lambda = \dfrac{\sum_{i=1}^{k} \alpha_i}{n + \sum_{i=1}^{k} \alpha_i} < 1$

as n increases $\underline{\lambda \to 0}$ it gets closer to

MLE

3a we have

$$\text{Rating} = \frac{nR}{n+m} + \frac{mC}{n+m}$$

where R → actual average rating of movie.

which can also be said as "Maximum Likelihood Estimation" or MLE

Hence MLE = R

n → number of votes for the movie

m → minimum number of votes required 2500

C = 5.5 (Mean rating of all report)

By above expression it could be worked out that

C is the estimate of Drichlet prior

$$C = \frac{\alpha_i}{\Sigma \alpha_i}$$

So for the scenario

$$\Sigma \alpha_i = 2500$$

$\alpha = 2500 \times 5.5$

$= \underline{13750}$

Top 10 movies

Imdb_id	Rating
1 → tt 5074352	8.345830
2 → tt 8108 8198	8.266026
3 → tt 8291224	8.207242
4 → tt 454470	8.108951
5 → tt 4430212	8.085411
6 → tt 3322420	8.085018
7 → tt 2356180	8.066135
8 → tt 0073707	8.061031
9 → tt 2283748	8.055511
10 → tt 1639426	8.039542