Pareto distribution $p(x)$ → Real life rules

γ distribution, β distribution

↳ 80-20 rule

a lot of problems why pareto ⇒ you can model distributions

Beta distribution

$$\text{Beta}(\mu|a,b) = \frac{\Gamma(a+b)}{\Gamma(a)\Gamma(b)} \mu^{a-1} (1-\mu)^{b-1}$$

μ → value of random variable

$$\Gamma(x) = \int_0^x u^{x-1} e^{-u} \, du$$

$$E(\mu) = \frac{a}{a+b}$$

$$V(\mu) = \frac{ab}{(a+b)^2 (a+b+1)}$$

Bayesian Bernoulli:

MLE : $\mu = \frac{m}{n}$ for m heads in n tosses

This is an overfit : if $n = m$ it will estimate $\mu = 1$ for all subsequent n

$[\cdot \cdot \cdot \cdot \cdot \cdot \cdot]$

$\mu = ??$

$$P(\mu/D) = \frac{P(D|\mu) \cdot P(\mu)}{P(D)}$$

$$\propto P(D/\mu)\, P(\mu)$$

$$\propto \prod_{n=1}^{n} \mu^{m} (1-\mu)^{n-m} \qquad p(\mu)$$

$$\propto \prod_{n=1}$$

next page

$$p(\mu/D) = \mathrm{Beta}(m+a, l+b)$$
$$= \mathrm{Beta}(a, b)$$

Beta$(\mu|a,b)$

$$p(\mu) = \frac{\Gamma(a+b)}{\Gamma(a)\Gamma(b)} \mu^{a-1} (1-\mu)^{b-1}$$

↳ prior probability

→ μ – how biased the coin is?

→ μ itself acts a random variable

→ Let a blacksmith want to model this [$p(\text{head}) = \mu$] for a coin

Lets take the distribution of μ as Beta distribution

$$p(\mu|D) = \frac{p(D/\mu) \cdot p(\mu)}{p(D)}$$

(having done the $\sqrt{}$ expts)

$$\propto p(D/\mu) \cdot p(\mu)$$

$$= p(x_1, x_2 \cdots\cdots x_n/\mu)\, p(\mu)$$

$$= \prod_{n=1}^{N} \left(p(x_n/\mu) \right) p(\mu)$$

$$= \prod_{n=1}^{N} \left(\mu^{x_n} (1-\mu)^{1-x_n} \right) p(\mu)$$

$$= \left[\mu^{m} (1-\mu)^{n-m} \frac{\Gamma(a+b)}{\Gamma(a)\Gamma(b)} \mu^{a-1} (1-\mu)^{b-1} \right]$$

$$\propto \mu^{m+a-1} (1-\mu)^{n-m+b-1}$$

$$\propto \mu^{m+a-1} (1-\mu)^{l+b-1}$$

$$p(\mu/D) \propto \mu^{m+a-1} (1-\mu)^{l+b-1} \qquad l = n-m$$

$$p(\mu/D) = \frac{\mu^{m+a-1}(1-\mu)^{l+b-1}}{\underbrace{\int \mu^{m+a-1}(1-\mu)^{l+b-1}\, d\mu}_{\text{normalization}}}$$

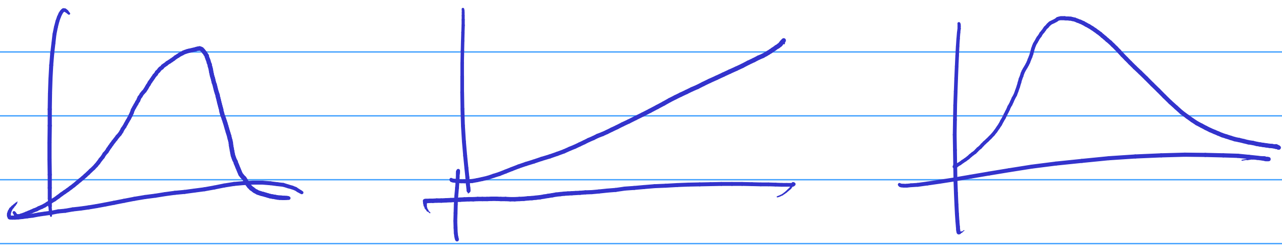
$$p(\mu/D) = \frac{\Gamma(m+a+l+b)}{\Gamma(m+a)\Gamma(l+b)} \mu^{m+a-1} (1-\mu)^{l+b-1}$$

$$p(\mu/D) = Beta(m+a, l+b)$$

posterior

$$p(\mu) = Beta(a, b)$$

$\hookrightarrow Beta(m+a, l+b)$



$$p(head) = E(\mu)$$

$$= \frac{m+a}{m+a+l+b}$$

$$MLE = \frac{m}{N} = \mu = \frac{m}{m+l}$$