Beta distribution

$$\mathrm{Beta}(\mu|a,b) = \frac{\Gamma(a+b)}{\Gamma(a)\cdot\Gamma(b)} \mu^{a-1} (1-\mu)^{b-1}, \quad \mu \in [0,1], \; a,b > 0$$

$$E[\mu] = \frac{a}{a+b}, \quad \mathrm{var}[\mu] = \frac{ab}{(a+b)^2(a+b+1)}$$

$\Gamma(x) = (x-1)!$

$\Gamma(x) = \int_0^\infty t^{x-1} \cdot e^{-t} \, dt$

Recall: $\mathrm{Bern}(x|\mu) = \mu^x (1-\mu)^{1-x}$

$$\underbrace{B(a,b) \triangleq \frac{\Gamma(a)\cdot\Gamma(b)}{\Gamma(a+b)}}_{\text{beta function}}$$

$$p(\mu|D) = \frac{p(D|\mu)\,p(\mu)}{\int p(D|\mu)\,p(\mu)\,d\mu}$$

$$= \frac{\prod_{n=1}^{N} p(x_n|\mu)\,p(\mu|a,b)}{\int_0^1 \prod_{n=1}^{N} p(x_n|\mu)\,p(\mu|a,b)\,d\mu}$$

$$= \frac{\mu^{(\#1)}(1-\mu)^{(\#0)} \cdot \frac{1}{B(a,b)} \cdot \mu^{a-1}(1-\mu)^{b-1}}{\int_0^1 \mu^{(\#1)}(1-\mu)^{(\#0)} \cdot \frac{1}{B(a,b)} \cdot \mu^{a-1}(1-\mu)^{b-1}\,d\mu}$$

$$= \frac{\mu^{(\#1)+a-1}(1-\mu)^{(\#0)+b-1}}{\underbrace{B((\#1)+a,\,(\#0)+b)}_{\int_0^1 \mu^{(\#1)+a-1}(1-\mu)^{(\#0)+b-1}\,d\mu}}$$

$$\int_0^1 x^a (1-x)^b \, dx = B(a+1, b+1)$$

Eulerian integral of first kind

$$= \frac{\mu^{(\#1)+a-1}(1-\mu)^{(\#0)+b-1}}{B((\#1)+a,(\#0)+b)} = \mathrm{Beta}(\mu\,|\,(\#1)+a,(\#0)+b)$$

Mean:

$$E[\mu] = \frac{(\#1)+a}{(\#1)+(\#0)+a+b}.$$

$D = \{0,1,1,0,1,0,1,1,1,0\}$

$\#0 = 4,\ \#1 = 6$

suppose $a=b=2$: $\frac{6+2}{10+4} = 0.57$

Maximum likelihood gives $\mu_{ML} = 0.6 = \frac{\#1}{\#0+\#1}$

Literature: Gregor Heinrich, Parameter estimation for text analysis. Technical Report, 2009.