- Let x be a hypothesized prob. of heads for a biased coin
- Let n be the number of coin flips
- Let h = # heads

  $t = n-h$ = # tails

The "data", D, is the # heads & # tails "observed"

$$\Pr(D|x) = \text{bin}(h;\, n, x)$$

$$= \binom{n}{h} x^h (1-x)^{n-h}$$

Max likelihood $\max_x \Pr(D|x)$

let $f(x) = \binom{n}{h} x^h (1-x)^{n-h}$

note: x which maximizes f also maximizes $\ln f$ ...

$$g(x) = \ln\left[\binom{n}{h} x^h (1-x)^{n-h}\right] = \ln\binom{n}{h} + h\ln x + (n-h)\ln(1-x)$$

$$g'(x) = \frac{h}{x} + \frac{n-h}{1-x}(-1)$$

$$g'(x)=0 \iff \frac{h}{x} = \frac{n-h}{1-x} \iff \frac{x}{h} = \frac{1-x}{n-h}$$

$$\iff (n-h)x = h(1-x)$$

$$\iff (n-h)x = h - hx$$

$$\iff nx = h$$

$$\iff \boxed{x = h/n}$$

· <u>Dist</u> $\int_0^1 f(x)\,dx = \binom{n}{h} \int_0^1 x^h (1-x)^{n-h}\, dx$

$= \binom{n}{h} \ \ldots$

Let $h = 0$ (no heads)

$$\int_0^1 f(x)\,dx = \int_0^1 (1-x)^n\, dx$$

$$= \left. \frac{-(1-x)^{n+1}}{n+1} \right|_0^1 = \frac{1}{n+1}$$

Thus, $\hat{f}(x) = (n+1)(1-x)^n$ is a dist. on $[0,1]$

$$E[\hat{f}] = \int_0^1 x \hat{f}(x)\, dx$$

$$= (n+1) \int_0^1 x (1-x)^n\, dx$$

$u = x \qquad dv = (1-x)^n dx$

$du = dx \qquad v = \frac{-(1-x)^{n+1}}{n+1}$

$$= (n+1) \left[\frac{-x(1-x)^{n+1}}{n+1} - \int \frac{-(1-x)^{n+1}}{n+1}\, dx \right]_0^1$$

$$= (n+1) \left[\frac{-x(1-x)^{n+1}}{n+1} - \frac{(1-x)^{n+2}}{(n+1)(n+2)} \right]_0^1$$

$$= \frac{(n+1)}{(n+1)(n+2)} = \boxed{\frac{1}{n+2}}$$

/

Javed Aslam

See $\boxed{8/14/98}$ notes for setup ...

- Bayesian / MAP approach :

prior probabilities

Bayes law : $\Pr(A|B) = \dfrac{\Pr(B|A)\cdot \Pr(A)}{\Pr(B)} = \dfrac{\Pr(B|A)\cdot\Pr(A)}{\sum_a \Pr(B|a)\cdot\Pr(a)}$

probability

- this is for a finite hypothesis space

- the analog for an infinite hypothesis space is, I believe;

prior density

See Barr & Zehna, → "Probability: Modeling Uncertainty" pp. 214-215

$$f(A|B) = \frac{\Pr(B|A)\cdot g(A)}{\int_a \Pr(B|a)\, g(a)\, da}$$

density

Our case:

our case - uniform prior, $g(x) = 1$ $\forall x \in [0,1]$

$$f(x|D) = \frac{\Pr(D|x)\cdot g(x)}{\cancel{\int \Pr(D|x)\cdot g(x)\,dx}}$$

$$= \frac{\binom{n}{h} x^h (1-x)^{n-h}}{\int_0^1 \binom{n}{h} x^h (1-x)^{n-h}\,dx}$$

$$= \frac{\binom{n}{h} x^h (1-x)^{n-h}}{\binom{n}{h}\cdot \frac{1}{(n+1)\binom{n}{h}}}$$ (see other 8/28/98 result)

$$= (n+1)\binom{n}{h} x^h (1-x)^{n-h}$$

Jamal Adam

Now, $f(x|D)$ is a probability density over hypotheses given an observed data set. What is the expected hypothesis?

Bayes point estimator: prob that next coin flip is heads

$$\int_0^1 x\, f(x|D)\, dx = (n+1)\binom{n}{h} \int_0^1 x^{h+1}(1-x)^{n-h}\, dx$$

$$= (n+1)\binom{n}{h} \frac{1}{(n+2)\binom{n+1}{h+1}}$$

$$= (n+1)\binom{n}{h} \frac{1}{(n+2)\left(\frac{n+1}{h+1}\right)\binom{n}{h}}$$

$$= \boxed{\frac{h+1}{n+2}}$$

Thus, $\frac{h+1}{n+2}$ is the expected or mean hypothesis using the MAP/Bayesian density over the hypothesis space.

Calculate bias of $\frac{h+1}{n+2}$ hypothesis.

- n - # trials
- p - prob. of success

$$\text{bias} = E\left[\frac{h+1}{n+2} - p\right]$$

$$= \sum_{h=0}^{n} \frac{h+1}{n+2} \binom{n}{h} p^h (1-p)^{n-h} \quad - \quad p$$

$$= \frac{1}{n+2}\left[\sum_{h=1}^{n} h\binom{n}{h} p^h (1-p)^{n-h} + \sum_{h=0}^{n} \binom{n}{h} p^h (1-p)^{n-h}\right] - p$$

$$= \frac{1}{n+2}\left[n\sum_{h=1}^{n} \binom{n-1}{h-1} p^h (1-p)^{n-h} + 1\right] - p$$

$$= \frac{1}{n+2}\left[n\sum_{k=0}^{n-1} \binom{n-1}{k} p^{k+1} (1-p)^{n-k-1} + 1\right] - p$$

$$= \frac{1}{n+2}\left[np\sum_{k=0}^{n-1} \binom{n-1}{k} p^{k} (1-p)^{(n-1)-k} + 1\right] - p$$

$$= \frac{1}{n+2}\left[np + 1\right] - p$$

$$= \frac{np+1}{n+2} - p$$

(Easier derivation: $E\left[\frac{h+1}{n+2} - p\right] = E\left[\frac{h+1}{n+2}\right] - E[p]$

$= \frac{1}{n+2} E[h+1] - p$

$= \frac{1}{n+2}\left(E[h] + E[1]\right) - p$

$= \frac{1}{n+2}(np+1) - p$ ✓)

$$= \frac{np+1-(np+2p)}{n+2}$$

$$= \frac{1-2p}{n+2}$$

- no bias at $1/2$; max bias at $p=0$ & $p=1$ ($\pm\frac{1}{n+2}$ bias)

Jared Ne

①

MSE of $\frac{h+1}{n+2}$ vs. $\frac{h}{n}$

1) $\frac{h+1}{n+2}$: $E\left[\left(\frac{h+1}{n+2}-p\right)^2\right] = E\left[\left(\frac{h+1-np-2p}{n+2}\right)^2\right]$

$$= \frac{1}{(n+2)^2} E\left[\left((h-np)+(1-2p)\right)^2\right]$$

$$= \frac{1}{(n+2)^2} E\left[(h-np)^2 + 2(h-np)(1-2p) + (1-2p)^2\right]$$

$$= \frac{1}{(n+2)^2}\left[E\left[(h-np)^2\right] + 2(1-2p)\,E\left[\cancel{(h-np)}^{\,0}\right] + (1-2p)^2\right]$$

$$= \frac{Var(h) + (1-2p)^2}{(n+2)^2}$$

$$= \boxed{\frac{np(1-p) + (1-2p)^2}{(n+2)^2}}$$

2) $\frac{h}{n}$: $E\left[\left(\frac{h}{n}-p\right)^2\right] = E\left[\left(\frac{h-np}{n}\right)^2\right]$

$$= \frac{1}{n^2} E\left((h-np)^2\right)$$

$$= \frac{Var(h)}{n^2}$$

$$= \boxed{\frac{np(1-p)}{n^2} = \frac{p(1-p)}{n}}$$

[redacted]

$\boxed{\int x^a (1-x)^b \, dx}$ 8/28/98 (1

· let $f_x(a,b) = x^a (1-x)^b$

$$\int f_x(a,b)\,dx = \int x^a (1-x)^b \, dx$$

$$= \int x^a \, d\left(\frac{(1-x)^{b+1}}{-(b+1)}\right)$$

$$= -\frac{x^a (1-x)^{b+1}}{b+1} + \frac{a}{b+1} \int x^{a-1} (1-x)^{b+1} \, dx$$

$$= -\frac{x^a (1-x)^{b+1}}{b+1} + \frac{a}{b+1} \int f_x(a-1, b+1)\, dx$$

$$= -\frac{x^a (1-x)^{b+1}}{b+1} + \frac{a}{b+1}\left[-\frac{x^{a-1}(1-x)^{b+2}}{b+2} + \frac{a-1}{b+2} \int f_x(a-2, b+2)\, dx \right]$$

$$= -\frac{x^a (1-x)^{b+1}}{b+1} - \frac{a\, x^{a-1} (1-x)^{b+2}}{(b+1)(b+2)} + \frac{a(a-1)}{(b+1)(b+2)} \int f_x(a-2, b+2)\, dx$$

$$\vdots$$

$$= -\frac{1}{b+1} x^a (1-x)^{b+1} - \frac{a}{(b+1)(b+2)} x^{a-1} (1-x)^{b+2} - \frac{a(a-1)}{(b+1)(b+2)(b+3)} x^{a-2} (1-x)^{b+3} \cdots - \frac{a(a-1)\cdots 2}{(b+1)(b+2)\cdots(b+a)} x (1-x)^{b+\cdots}$$

$$+ \frac{a(a-1)\cdots 1}{(b+1)(b+2)\cdots(b+a)} \underbrace{\int f_x(0, b+a)\, dx}_{\frac{(1-x)^{b+a+1}}{-(b+a+1)}}$$

$$\boxed{= -\sum_{i=0}^{a} \left[\prod_{j=1}^{i} \left(\frac{a-j+1}{b+j}\right) \cdot \frac{x^{a-i} (1-x)^{b+i+1}}{b+i+1} \right]}$$

<u>Note</u>: $\boxed{\int_0^1 x^a (1-x)^b \, dx = \frac{a\cdot(a-1)\cdots 1}{(b+1)(b+2)\cdots(b+a+1)} = \frac{a!\, b!}{(b+a+1)!} = \frac{1}{(b+a+1)\binom{a+b}{a}}}$