(1) Bishop 1.3

a. $p(\text{apple}) = p(a|r)p(r) + p(a|b)p(b) + p(a|g)p(g)$

$= 0.2 \times 0.3 + 0.2 \times 0.5 + 0.6 \times 0.3$

$= 0.06 + 0.1 + 0.18$

$= \boxed{0.34}$

b. $p(g|\text{orange}) = \dfrac{P(\text{orange}|g)P(g)}{P(\text{orange})} = \dfrac{0.6 \times 0.3}{0.4 \times 0.2 + 0.2 \times 0.5 + 0.6 \times 0.3}$

$= \dfrac{0.18}{0.08 + 0.1 + 0.18}$

$= \boxed{0.5}$

(2) Bishop 2.1

$\text{Bernoulli}(x|\mu) = \mu^x (1-\mu)^{1-x}$

(2.257) $\Sigma_0^1 p(x|\mu) = \mu^0 (1-\mu)^1 + \mu^1 (1-\mu)^0 = (1-\mu) + \mu = 1$ ✓

(2.258) $E[x] = \sum_{x=0}^{1} x\, p(x|\mu) = 0 \cdot \mu^0 (1-\mu)^1 + 1 \cdot \mu^1 (1-\mu)^0 = \mu$ ✓

(2.259) $\text{var}[x] = E[x^2] - E[x]^2$

$= \sum_{x=0}^{1} x^2 p(x|\mu) - E[x]^2$

$= \mu - \mu^2$

$= \mu(1-\mu)$ ✓

(3) Bishop 2.6

(2.265) $\int_0^1 \mu^{a-1} (1-\mu)^{b-1} d\mu = \dfrac{\Gamma(a)\Gamma(b)}{\Gamma(a+b)}$

(2.267) $E[\mu] = \dfrac{\Gamma(a+b)}{\Gamma(a)\Gamma(b)} \int_0^1 \mu^a (1-\mu)^{b-1} d\mu = \dfrac{\Gamma(a+b)}{\Gamma(a)\Gamma(b)} \cdot \dfrac{\Gamma(a+1)\Gamma(b)}{\Gamma(a+b+1)}$

$= \dfrac{a}{a+b}$ ✓

(2.268) $\text{var}[\mu] = \frac{\Gamma(a+b)}{\Gamma(a)\Gamma(b)} \int_0^1 \mu^{a+1}(1-\mu)^{b-1} - E[\mu]^2$

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

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

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

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

(2.269) $\frac{df(x)}{dx} = 0$

$$\frac{\Gamma(a+b)}{\Gamma(a)\Gamma(b)}\left((a-1)\mu^{a-2}(1-\mu)^{b-1} - (b-1)\mu^{a-1}(1-\mu)^{b-2}\right) = 0$$

$$\mu^{(a-2)}(1-\mu)^{b-2}[(a-1)(1-\mu)-(b-1)\mu] = 0$$

$$a-1-a\mu+\mu-b\mu+\mu = 0$$

$$\mu = \frac{a-1}{a+b-2} \checkmark$$

(4) (a) posterior $\propto$ likelihood $\times$ prior

$$\Rightarrow \sum_{i=1}^{N} \mu^{x^{(i)}}(1-\mu)^{(1-x^{(i)})} \times \frac{\Gamma(a+b)}{\Gamma(a)\Gamma(b)}\mu^{a-1}(1-\mu)^{b-1}$$

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

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

where $m = \sum_{i=1}^N x^{(i)}$ & $l = \sum_{i=1}^N (1-x^{(i)})$

posterior is related to prior, with extra m heads & l tails, thus m should be added to a & l should be added to b.

And $p(\mu|D) = p(\mu; m, l, a, b) \propto \mu^{m+a-1}(1-\mu)^{l+b-1}$

(b) (2.19) $P(x=1|D) = \int_0^1 p(x=1|\mu)P(\mu|D)d\mu = \int_0^1 \mu p(\mu|D)d\mu = E[\mu|D]$

by (2.18) $\Rightarrow p(\mu|m,l,a,b) = \frac{\Gamma(m+a+l+b)}{\Gamma(m+a)\Gamma(l+b)}\mu^{m+a-1}(1-\mu)^{l+b-1}$

$$p(x=1|D) = \int_0^1 \mu \cdot \frac{\Gamma(m+a+l+b)}{\Gamma(m+a)\Gamma(l+b)}\mu^{m+a-1}(1-\mu)^{l+b-1}$$

$$= \frac{\Gamma(m+a+l+b)}{\Gamma(m+a)\Gamma(l+b)} \int_0^1 \mu^{m+a}(1-\mu)^{l+b-1}$$

by (2.265) $\int_0^1 \mu^{a-1}(1-\mu)^{b-1} d\mu = \frac{\Gamma(a)\Gamma(b)}{\Gamma(a+b)}$

$$P(X=1|D) = \frac{\Gamma(m+a+l+b)}{\Gamma(m+a)\Gamma(l+b)} \cdot \frac{\Gamma(m+a+1)\Gamma(l+b)}{\Gamma(m+a+1+l+b)} = \frac{m+a}{m+a+l+b} \quad \checkmark$$

(5) Bishop 1.11

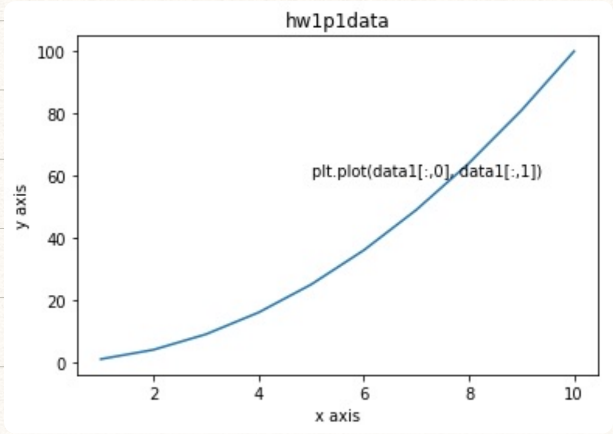
$$\ln p(X|\mu,\sigma^2) = -\frac{1}{2\sigma^2}\sum_{n=1}^{N}(x_n-\mu)^2 - \frac{N}{2}\ln\sigma^2 - \frac{N}{2}\ln(2\pi)$$

$$\frac{d}{d\mu}\ln p(X|\mu,\sigma^2) = \frac{1}{\sigma^2}\sum_{n=1}^{N}(x_n-\mu) = 0$$

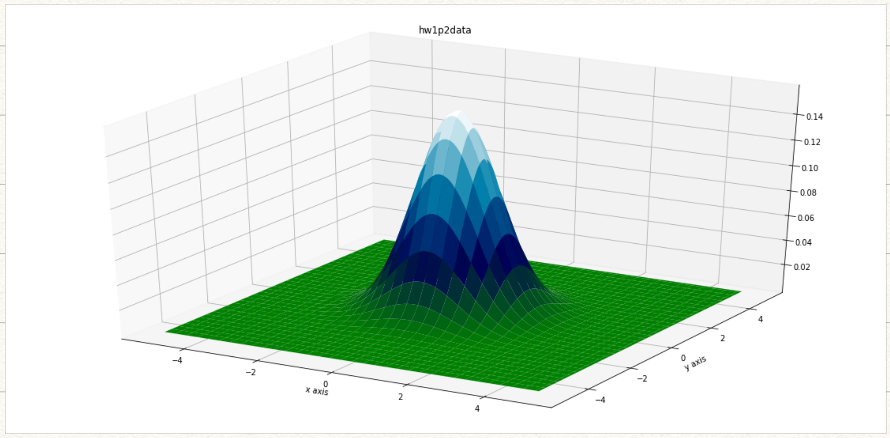
$$\sum_{n=1}^{N} x_n = n\mu$$

$$(1.55) \quad \mu = \frac{1}{N}\sum_{n=1}^{N} x_n \quad \checkmark$$

(6)



(7)



(bonus)

$$P(\text{last}b \mid \text{room}g, \text{letter}b) = \frac{P(\text{room}g, \text{letter}b \mid \text{last}b)\, P(\text{last}b)}{P(\text{room}g, \text{letter}b)}$$

$$P(\text{last}g \mid \text{room}g, \text{letter}b) = \frac{P(\text{room}g, \text{letter}b \mid \text{last}g)\, P(\text{last}g)}{P(\text{room}g, \text{letter}b)}$$

$$P(\text{room}g, \text{letter}b \mid \text{last}b) = P(\text{room}g \mid \text{last}b) \times P(\text{letter}b \mid \text{room}g, \text{last}b)$$
$$= \frac{1}{3} \times \frac{2}{3}$$

$$P(\text{last}b) = \frac{1}{2}$$

$$\Rightarrow P(\text{last}b \mid \text{room}g, \text{letter}b) = \frac{\frac{1}{3} \times \frac{2}{3} \times \frac{1}{2}}{P(\text{room}g, \text{letter}b)} = \frac{\frac{1}{9}}{P(\text{room}g, \text{letter}b)}$$

$$P(\text{room}g, \text{letter}b \mid \text{last}g) = P(\text{room}g \mid \text{last}g) \times P(\text{letter}b \mid \text{room}g, \text{last}g)$$
$$= \frac{1}{2} \times \frac{1}{3}$$

$$P(\text{last}g) = \frac{1}{2}$$

$$\Rightarrow P(\text{last}g \mid \text{room}g, \text{letter}b) = \frac{\frac{1}{2} \times \frac{1}{3} \times \frac{1}{2}}{P(\text{room}g, \text{letter}b)} = \frac{\frac{1}{12}}{P(\text{room}g, \text{letter}b)}$$

$$\text{boy} : \text{girl} = \frac{1}{9} : \frac{1}{12} = 4 : 3$$