HMM Problem (Jason Eisner's ice creams).

(i) likelihood (forward algorithm)

In each entry we compute $\alpha_t(j) = Pr(O_{1:t}, q_t = j | \lambda)$

$$= \sum_i Pr(O_{1:t-1}, q_{t-1} = i | \lambda) \cdot Pr(q_t = j | q_{t-1} = i) \cdot Pr(O_t | q_t = j)$$

$$= \sum_i \alpha_{t-1}(i) A_{ij} B_j(O_t)$$

| | 1 | 3 | 3 |
|---|---|---|---|
| HOT | $0.8\times0.2=0.16$ | $0.16\times0.6\times0.4+$ $0.1\times0.5\times0.4=0.0584$ | $0.0584\times0.6\times0.4$ $+0.0114\times0.5\times0.4=0.0163$ |
| COLD | $0.2\times0.5=0.1$ | $0.16\times0.4\times0.1$ $+0.1\times0.5\times0.1=0.0114$ | $0.0584\times0.4\times0.1$ $+0.0114\times0.5\times0.1=0.002906$ |

$$\Rightarrow Pr(\text{"1 3 3"} | \lambda) = Pr(\text{"1 3 3"}, q_3 = hot) + Pr(\text{"1 3 3"}, q_3 = cold)$$

$$= 0.0163 + 0.002906 = 0.0192.$$

We can also use backward algorithm to compute $Pr(\text{"133"}|\lambda)$.

In each entry we compute:

$$\beta_t(j) = Pr(O_{t+1:T}, | q_t = j, \lambda)$$

$$= \sum_i Pr(O_{t+2:T} | q_{t+1} = i, \lambda) \cdot Pr(q_{t+1} = i | q_t = j) \cdot Pr(O_{t+1} | q_{t+1} = i)$$

$$= \sum_i \beta_{t+1}(i) A_{ji} B_i(O_{t+1})$$

| | 1 | 3 | 3 |
|---|---|---|---|
| HOT | $(0.28\times0.6\times0.1)+$ $(0.25\times0.4\times0.1)=0.0772$ | $(1\times0.6\times0.4)+$ $(1\times0.4\times0.1)=0.28$ | 1 |
| COLD | $(0.28\times0.5\times0.4)+$ $(0.25\times0.5\times0.1)=0.0685$ | $(1\times0.5\times0.4)+$ $(1\times0.5\times0.1)=0.25$ | 1 |

$$\Rightarrow P(\text{"133"}|\lambda) = \pi(hot)\cdot Pr(\text{"133"}|q_1 = hot, \lambda) + \pi(cold)\cdot Pr(\text{"133"}|q_1 = cold, \lambda)$$

$$= (0.0772\times0.8\times0.2) + (0.0685\times0.2\times0.5)$$

$= 0.0192$ (same result as we got using forward algorithm).

(ii) Decoding: we'll apply Viterbi algorithm to compute:

$$V_t(j) = \max_{q_{1:t-1}} Pr(q_{1:t-1}, O_{1:t}, q_t = j \mid \lambda)$$

$$= \max_i \left\{ \max_{q_{1:t-2}} Pr(q_{1:t-2}, O_{1:t-1}, q_{t-1} = i \mid \lambda) \cdot Pr(q_t = j \mid q_{t-1} = i) \cdot Pr(O_t \mid q_t = j) \right\}$$

$$= \max_i \left\{ V_{t-1}(i) A_{ij} B_j(O_t) \right\}.$$

| | 1 | 3 | 3 |
|---|---|---|---|
| HOT | 0.8×0.2=0.16 | max{0.16×0.6×0.4, 0.1×0.5×0.4} =0.0384 | max{0.0384×0.6×0.4, 0.0064×0.5×0.4} =0.009216 |
| COLD | 0.2×0.5=0.1 | max{0.16×0.4×0.1, 0.1×0.5×0.1}=0.0064 | max{0.0384×0.4×0.1, 0.0064×0.5×0.1} =0.001536 |

By tracing back those maximal values we got:

$$\max_{q_{1:3}} Pr(\text{"1 3 3"}, q_{1:3} \mid \lambda) = \text{"hot hot hot"}.$$

(iii) Learning: the Baum-Welch algorithm is quite complicated, and I shall only show the computation for 1 E-step and 1 M-step for $A_{hh}$ and $B$. I'll use parameters in (i)&(ii) as initial parameters: $\lambda_0 = \{\pi, A, B\}$

① E-step:

- compute $\xi_t(1,1)$: since we've computed $\alpha, \beta$ in (i), we could then easily obtain all $\xi_t(i,j)$:

$$\xi_t(i,j) = Pr(q_t = i, q_{t+1} = j, \vec{O} \mid \lambda_0)$$

$$= Pr(O_{1:t}, q_t = i \mid \lambda_0) \cdot Pr(O_{t+2:T} \mid q_{t+1} = j, \lambda_0)$$

$$\cdot Pr(O_t \mid q_t = i, \lambda_0) \cdot Pr(q_{t+1} = j \mid q_t = i)$$

$$= \alpha_t(i) \beta_{t+1}(j) \cdot A_{ij} B_j(O_t)$$

$$\Rightarrow \xi_1(h,h) = \alpha_1(h) \beta_2(h) A_{hh} B_h(1)$$

$$= 0.16 \times 0.28 \times 0.6 \times 0.2 = 0.005376$$

$$\xi_2(h,h) = \alpha_2(h) \beta_3(h) A_{hh} B_h(3)$$

$$= 0.0584 \times 1 \times 0.6 \times 0.4 = 0.014016$$

- compute $\gamma_t(h)$:

$$\gamma_1(h) = \alpha_1(h)\,\beta_1(h) = 0.16 \times 0.0772 = 0.12352$$

$$\gamma_2(h) = \alpha_2(h)\,\beta_2(h) = 0.0584 \times 0.28 = 0.016352$$

$$\gamma_3(h) = \alpha_3(h)\,\beta_3(h) = 0.0163 \times 1 = 0.0163$$

- Now, update $A_{hh}$ and $B_{h1}$:

$$A_{hh} = \frac{\xi_1(hh) + \xi_2(hh)}{\gamma_1(h) + \gamma_2(h)} = 0.1386$$

$$B_{h1} = \frac{\gamma_1(h)}{\gamma_1(h) + \gamma_2(h) + \gamma_3(h)} = 0.7909$$

SPP Problem.

Let $U(\alpha,\beta)$ denote the expected utility (return) if the player pays $\alpha$ dollars at first, with the coin having probability $\beta$ to show its head in each toss:

$$U(\alpha,\beta) = \sum_{k=1}^{\infty}\left(\sqrt{\alpha\cdot 2^k}-\alpha\right)\cdot\beta^k$$

$$= \sqrt{\alpha}\left(\frac{\sqrt{2}\beta}{1-\sqrt{2}\beta}\right)-\alpha$$

To obtain maximum return, setting $\frac{\partial U}{\partial\alpha}=0$ we have:

$\alpha^*(\beta) = \frac{\beta^2}{2(1-\sqrt{2}\beta)^2}$, and the optimal utility would be:

$$U(\alpha^*(\beta),\beta) = \frac{\beta^2}{2(1-\sqrt{2}\beta)^2}$$

For a fair coin, by taking $\beta=\frac{1}{2}$ we have:

$$U(\alpha^*(\tfrac{1}{2}),\tfrac{1}{2}) = \frac{3+2\sqrt{2}}{4} \approx 1.457.$$

$$\alpha^*(\tfrac{1}{2}) = \frac{3+2\sqrt{2}}{4} \approx 1.457.$$

---

As we can see, $U(\alpha^*(\beta),\beta) = \frac{\beta^2}{2(1-\sqrt{2}\beta)^2} > 0$, so a rational player would always have positive expected return, and the bookies will lose in expectation (since this is a zero-sum game). Now suppose the bookies covertly use a coin with $\alpha\beta < \frac{1}{2}$, but they shall still claim that the coin is fair. Hence, all players, if acting rationally, would still pay $\alpha^*(\frac{1}{2}) = 1.457$ dollars at each round. Their expected utility would then actually be:

$$U(\alpha^*(\tfrac{1}{2}), \beta) = \sqrt{\alpha^*(\tfrac{1}{2})}\left(\frac{\sqrt{2}\beta}{1-\sqrt{2}\beta}\right) - \alpha^*(\tfrac{1}{2}).$$

Setting $U(\alpha^*(\frac{1}{2}), \beta) = 0$ we have:

$$\beta_0 = \frac{\sqrt{\alpha^*(\frac{1}{2})}}{\sqrt{2}\left(1+\sqrt{\alpha^*(\frac{1}{2})}\right)} \approx 0.387.$$

Hence, if the bookmakers covertly set $\beta < \beta_0$, they'll have positive expected return (and players will lose in expectation).

---

Now if an info-broker sells the following info:

"the bookies are using a coin with $\beta = 0.3$ secretly"

the players would, of course, adjust their strategies (since if they still pay $\alpha^*(\frac{1}{2})$ dollars at each round, they'll have negative expected return). Taking $\beta = 0.3$, the optimal pay $\alpha^*(\beta)$ is:

$$\alpha^*(0.3) = 0.1358$$

$$\Rightarrow U(\alpha^*(0.3), \beta = 0.3) = \alpha^*(0.3) = 0.1358$$

Hence, the price for this information would be the different between: 1) the optimal return after the players know that the bookies are cheating; and 2) their expected return when they don't know this conspiracy:

If the players don't know the bookies are cheating, they'll pay $\alpha^*(\frac{1}{2})$ at each round, their expected return would be:

$$U(\alpha^*(\tfrac{1}{2}), 0.3) = \sqrt{\alpha^*(\tfrac{1}{2})}\left(\frac{\sqrt{2}\cdot 0.3}{1-\sqrt{2}\cdot 0.3}\right) - \alpha^*(\tfrac{1}{2})$$

$$\approx -0.568$$

Then the price would be:

$$p = U(\alpha^*(0.3), 0.3) - U(\alpha^*(\tfrac{1}{2}), 0.3)$$

$$\approx 0.7038.$$