# EE3731C Tutorial - Statistical Signal 2

## Department of Electrical and Computer Engineering

1. Consider

$$\begin{aligned}
p(x_1, x_2, x_3) &= p(x_1)p(x_2, x_3|x_1) \\
&= p(x_1)p(x_2|x_1)p(x_3|x_1, x_2) \\
&= p(x_1)p(x_2|x_1)p(x_3|x_2), \text{by conditional independence of } x_1 \text{ and } x_3 \text{ given } x_2
\end{aligned}$$

   In this case, $x_1 \to x_2 \to x_3$ are said to form a Markov Chain. Similarly,

$$\begin{aligned}
p(x_1, x_2, x_3) &= p(x_3)p(x_1, x_2|x_3) \\
&= p(x_3)p(x_2|x_3)p(x_1|x_2, x_3) \\
&= p(x_3)p(x_2|x_3)p(x_1|x_2), \text{by conditional independence of } x_1 \text{ and } x_3 \text{ given } x_2
\end{aligned}$$

   Therefore if $x_1 \to x_2 \to x_3$ is a Markov Chain, then $x_3 \to x_2 \to x_1$ is also a Markov Chain.

2. Let $x$ be the number of heads in 10 coin toss. Then $x$ is binomially distributed with number of trials equal 10 and probability of success equal to $q$

   (a)

$$\begin{aligned}
q_{ML} &= \underset{q}{\operatorname{argmax}}\, p(x = 9|q) \\
&= \underset{q}{\operatorname{argmax}} \binom{10}{9} q^9(1-q) \\
&= \underset{q}{\operatorname{argmax}}\, q^9 - q^{10} \qquad (1)
\end{aligned}$$

   Differentiating Eq. (1) with respect to $q$, we get

$$\frac{\partial}{\partial q}\left(q^9 - q^{10}\right) = 9q^8 - 10q^9$$

   Setting the above to 0, we get $q_{ML} = \frac{9}{10}$

(b) First let's compute $p(q|x=9)$:

$$\begin{aligned}
p(q|x=9) &= \frac{p(x=9|q)p(q)}{p(x=9)} \\
&= \frac{\binom{10}{9}q^9(1-q)2q}{\int_0^1 p(x=9|q)p(q)dq} \\
&= \frac{\binom{10}{9}q^9(1-q)2q}{\int_0^1 \binom{10}{9}q^9(1-q)2qdq} \\
&= \frac{q^{10}(1-q)}{\int_0^1 q^{10}-q^{11}dq} \\
&= \frac{q^{10}(1-q)}{\left[\frac{1}{11}q^{11}-\frac{1}{12}q^{12}\right]_0^1} \\
&= \frac{q^{10}(1-q)}{\left[\frac{1}{11}-\frac{1}{12}\right]} = 132q^{10}(1-q)
\end{aligned}$$

Therefore

$$\begin{aligned}
q_{MAP} &= \operatorname*{argmax}_q p(q|x=9) \\
&= \operatorname*{argmax}_q 132q^{10}(1-q) \\
&= \operatorname*{argmax}_q q^{10}(1-q) \qquad (2)
\end{aligned}$$

Differentiating Eq. (2) with respect to $q$, we get

$$\frac{\partial}{\partial q}\left(q^{10}-q^{11}\right) = 10q^9 - 11q^{10}$$

Setting the above to 0, we get $q_{MAP} = \frac{10}{11}$. Notice that the $q_{MAP} > q_{ML}$ because the prior is pulling the estimate of $q$ to be bigger because larger values of $q$ has a bigger prior probability.

(c)

$$\begin{aligned}
q_{MMSE} &= E_{p(q|x=9)}(q) \\
&= \int_0^1 qp(q|x=9)dq \\
&= \int_0^1 132q^{11}(1-q)dq \\
&= 132\int_0^1 q^{11}-q^{12}dq \\
&= 132\left[\frac{1}{12}q^{12}-\frac{1}{13}q^{13}\right]_0^1 \\
&= 132\left[\frac{1}{12}-\frac{1}{13}\right] = \frac{11}{13}
\end{aligned}$$

The posterior probability distribution (Figure 1) peaks at $\mu_{MAP} = 10/11$. $q_{MMSE}$ is equivalent to the mean of the posterior distribution. Because there is so much more area under the curve on the left of $\mu_{MAP}$ compared to the right of $\mu_{MAP}$, therefore, the mean of the posterior distribution (i.e., $q_{MMSE}$) is pulled towards the left and is smaller than $q_{MAP}$ (and even $q_{ML}$)

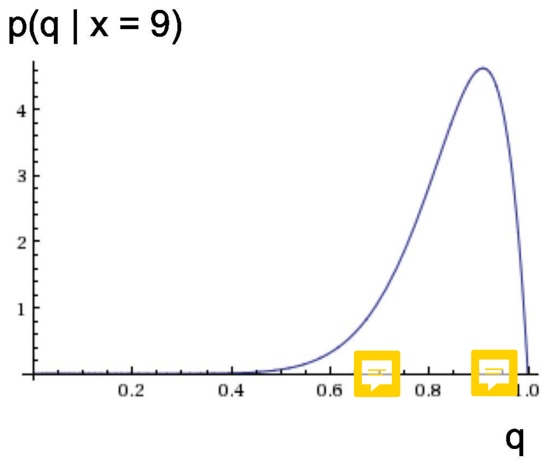


Figure 1: Posterior probability distribution: $p(q|x = 9)$

$$
\begin{aligned}
MSE(x = 9) &= E_{p(q|x=9)}(q - q_{MMSE})^2 \\
&= \int_0^1 p(q|x = 9)(q - \frac{11}{13})^2 dq \\
&= 132 \int_0^1 q^{10}(1 - q)(q - \frac{11}{13})^2 dq = \frac{11}{1183}
\end{aligned}
$$

3. (a) MMSE

   (b) From the last tutorial,

$$
p(x|y) = \begin{cases} \frac{x+y}{y+1/2} & \text{for } 0 \le x, y \le 1 \\ 0 & \text{otherwise} \end{cases}
$$

Therefore,

$$\begin{aligned}
x_n^{MMSE} = E(x_n|y_n) &= \int_0^1 xp(x_n|y_n)dx \\
&= \frac{1}{y+1/2}\int_0^1 x^2 + yxdx \\
&= \frac{1}{y+1/2}\left[\frac{x^3}{3} + \frac{yx^2}{2}\right]_0^1 \\
&= \frac{1}{y+1/2}\left(\frac{1}{3} + \frac{y}{2}\right)
\end{aligned}$$

Therefore average payout is

$$\int_0^1\int_0^1 p(x,y)(x_{MMSE}-x)^2 dxdy = \int_0^1\int_0^1 (x+y)\bigg(\frac{3y+2}{6y+3} - x\bigg)^2 dxdy$$

This is a complicated integral and according to wolfram alpha, the integral is equal to $\frac{1}{144}(12 - \log(3)) \approx 0.0757$. Let's try to compute this integral analytically. First recall that $x_{MMSE} = E(x|y)$, and so

$$\begin{aligned}
\int_0^1\int_0^1 p(x,y)(x_{MMSE}-x)^2 dxdy &= \int_0^1 p(y)\int_0^1 p(x|y)(x_{MMSE}-x)^2 dxdy \\
&= \int_0^1 p(y)\mathrm{Var}(x|y)dy
\end{aligned}$$

We get

$$\begin{aligned}
E(x^2|y) &= \frac{1}{y+1/2}\int_0^1 x^3 + yx^2 dx \\
&= \frac{1}{y+1/2}\left[\frac{1}{4}x^4 + \frac{1}{3}yx^3\right]_0^1 \\
&= \frac{1}{y+1/2}\left(\frac{1}{4} + \frac{1}{3}y\right)
\end{aligned}$$

Therefore

$$\begin{aligned}
\mathrm{Var}(x|y) &= E(x^2|y) - (E(x|y))^2 \\
&= \frac{1}{y+1/2}\left(\frac{1}{4} + \frac{1}{3}y\right) - \left(\frac{1}{y+1/2}\Big(\frac{1}{3} + \frac{y}{2}\Big)\right)^2
\end{aligned}$$

Recalling that $p(y) = y + 1/2$ for $y$ between 0 and 1, we get

$$\begin{aligned}
\int_0^1 p(y)\mathrm{Var}(x|y)dy &= \int_0^1 \frac{1}{4} + \frac{1}{3}y - \frac{1}{y+1/2}\left(\frac{1}{3} + \frac{y}{2}\right)^2 dy \\
&= \frac{1}{4} + \frac{1}{6} - \int_0^1 \frac{1}{y+1/2}\left(\frac{1}{3} + \frac{y}{2}\right)^2 dy \\
&= \frac{5}{12} - \int_{1/2}^{3/2} \frac{1}{z}\left(\frac{1}{3} + \frac{z-1/2}{2}\right)^2 dz \\
&= \frac{5}{12} - \int_{1/2}^{3/2} \frac{1}{z}\left(\frac{1+6z}{12}\right)^2 dz \\
&= \frac{5}{12} - \int_{1/2}^{3/2} \frac{1}{z}\left(\frac{1+12z+36z^2}{144}\right) dz \\
&= \frac{5}{12} - \frac{1}{144}\int_{1/2}^{3/2} \frac{1}{z} + 12 + 36z dz \\
&= \frac{5}{12} - \frac{1}{144}\left[\log z + 12z + 18z^2\right]_{1/2}^{3/2} \\
&= \frac{5}{12} - \frac{1}{144}(\log 3 + 48) \approx 0.0757
\end{aligned}$$

Therefore, the student pays me 0.0757 dollars on average each day.

(c) This is actually a trick question. The optimal strategy is technically the MAP strategy. However because $p(x|y)$ is a continuous distribution and so no matter what the student guesses, he or she has 0 probability of guessing it correctly. Therefore, I pay the student 0 dollars on average each day despite the student using an optimal MAP strategy.