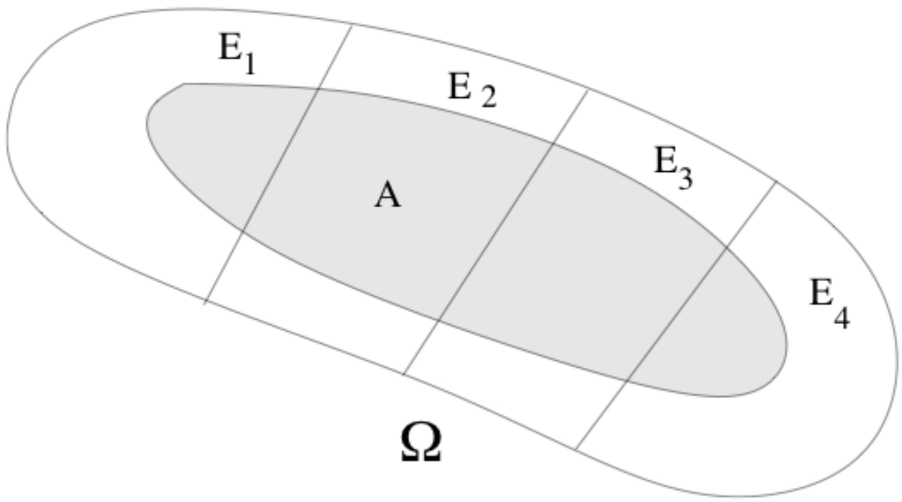
Lecture 12-14

2.10 The law of total probability and Bayes formula

Events $E_1, E_2, \cdots, E_k$ are said to form a partition of Ω if $E_1, \cdots, E_k$ are mutually disjoint and $E_1 \cup E_2 \cup \cdots \cup E_k = \Omega$



The law of total probability: For any event A and a partition $(E_1, E_2, \cdots, E_k)$ of Ω,

$$P(A) = P(AE_1) + P(AE_2) + \cdots + P(AE_k)$$
$$= P(A|E_1)P(E_1) + P(A|E_2)P(E_2) + \cdots + P(A|E_k)P(E_k)$$

Bayes's formula:

$$P(E_i|A) = \frac{P(AE_i)}{P(A)} = \frac{P(A|E_i)P(E_i)}{P(A)}$$
$$= \frac{P(A|E_i)P(E_i)}{P(A|E_1)P(E_1) + \cdots + P(A|E_k)P(E_k)}$$

The law of total probability

EX 2.10.2)

Step 1: Roll a die. Let X be the # shown in the die.

Step 2: Toss a coin X times. Let Y be the # of heads.

Q: $P\{Y=3\}$ and $P(X=3 \mid Y=3)$?

Define $E_i = \{X = i\}$ for $1 \le i \le 6$.

By the law of total probability,

$$P(Y=3) = \sum_{i=1}^{6} P(Y=3 \mid E_i)\, P(E_i)$$

$$P(Y=3 \mid E_i) = \begin{cases} 0 & \text{if } i \le 2 \\ \binom{i}{3}\left(\frac{1}{2}\right)^i & \text{if } i \ge 3 \end{cases} \qquad P(E_i) = \frac{1}{6}$$

Hence,
$$\begin{aligned} P(Y=3) &= \sum_{i=3}^{6} \binom{i}{3}\left(\frac{1}{2}\right)^i \cdot \frac{1}{6} \\ &= \frac{1}{6}\left(\binom{3}{3}\cdot 2^{-3} + \binom{4}{3}2^{-4} + \binom{5}{3}2^{-5} + \binom{6}{3}2^{-6}\right) \\ &= \frac{1}{6}\left(\frac{1}{8} + \frac{4}{16} + \frac{5\,4\,3}{3\,2\,1}\cdot\frac{1}{2^5} + \frac{6\,5\,4}{3\,2\,1}\cdot\frac{1}{2^6}\right) \\ &= \frac{1}{6}. \end{aligned}$$

$$P(E_3 \mid Y=3) = \frac{P(Y=3, E_3)}{P(Y)} = \frac{\frac{1}{6}\times\frac{1}{8}}{\frac{1}{6}} = \frac{1}{8}$$

Ex 2.10.4) Let $0<p<0.5$. Two biased coins.

First coin shows head with prob p.

Second coin shows head with prob. $q = 1-p$.

1. First choose a coin at random.
2. Toss the chosen coin n times

X = # of heads shown in the coin.

Q: Pmf, mean and variance of X?

Let A be the event that the first coin is chosen.

$$p_X(k) = P\{X=k\} = P(\{X=k\}\cap A) + P(\{X=k\}\cap A^c\}$$
$$= P(X=k\,|\,A)\,P(A) + P(X=k\,|\,A^c)\,P(A^c)$$
$$= \binom{n}{k} p^k (1-p)^{n-k}\cdot\frac{1}{2} + \binom{n}{k} q^k (1-q)^{n-k}\cdot\frac{1}{2}$$
$$= \frac{1}{2}\binom{n}{k}\left(p^k(1-p)^{n-k} + (1-p)^k p^{n-k}\right)$$

$$E[X] = \sum_k k\, p_X(k) = \sum_k k\, P\{X=k\}$$
$$= \sum_k k\left(P(X=k\,|\,A)P(A) + P(X=k\,|\,A^c)P(A^c)\right)$$
$$= \left(\sum_k k\, P(X=k\,|\,A)\right) P(A) + \left(\sum_k k\, P(X=k\,|\,A^c)\right) P(A^c)$$
$$= E(X\,|\,A)\,P(A) + E(X\,|\,A^c)\,P(A^c)$$
$$= np\cdot\frac{1}{2} + nq\frac{1}{2} = \frac{n}{2}$$

* The law of total probability $\Rightarrow$ Conditional expectation

For a partition $E_1, E_2, \cdots, E_k$ of Ω,

$$E[X] = \sum_{1=1}^{k} E[X\,|\,E_\hat{\imath}]\,P(E_\hat{\imath})$$

$$Var(X) = E[X^2] - E[X]^2$$

Let $Y = X^2$

$$E[Y] = E[X^2 | A]P(A) - E[X^2 | A^c]P(A^c)$$

$$= (Var(X|A) + E[X|A]^2)P(A) + (Var(X|A^c) + E[X|A^c]^2)P(A^c)$$

$$= (npq + n^2p^2)\frac{1}{2} + (npq + n^2q^2)\frac{1}{2}$$

$$= \frac{n^2(p^2+q^2)}{2} + npq$$

$$Var(X) = E[X^2] - E[X]^2 = \frac{n^2(p^2+q^2)}{2} + npq - \frac{n^2}{4}$$

$$= n^2\left(\frac{p^2+q^2}{2} - \frac{1}{4}\right) + npq.$$

2.11 Binary hypothesis testing

EX) A radar system launches a signal, receives a return signal with strength X.

H_0 : There is no enemy aircraft
H_1 : There is an enemy aircraft

After extensive testing using our airplanes, we found the following likelihood matrix.

	$X=1$	$X=2$	$X=3$	$X=4$
H_1	0.0	0.1	0.3	0.6
H_0	0.4	0.3	0.2	0.1

→ In the likelihood matrix sum of the numbers in each row = 1

A. Maximum likelihood hypothesis testing

Q1) Suppose you receives a radio signal with strength $X=3$. Which hypothesis maximizes the probability of $X=3$?
⇒ H_1.

Q2) If you send a signal, can you predict the distribution of X from the likelihood matrix?
⇒ No because we do not know $P(H_0)$ and $P(H_1)$.

Def) The Maximum Likelihood (ML) decision rule :
If $X=j$ is observed, choose the hypothesis H_i to maximize $P(X=j \mid H_i)$

	$X=1$	$X=2$	$X=3$	$X=4$
H_1	0.0	0.1	0.3	0.6
H_0	0.4	0.3	0.2	0.1

ML decisions

Likelihood ratio $\Lambda(k) = \dfrac{P(X=k \mid H_1)}{P(X=k \mid H_0)}$

miss alarm

	$X=1$	$X=2$	$X=3$	$X=4$
H_1	0.0	0.1	0.3	0.6
H_0	0.4	0.3	0.2	0.1

ML decisions

False alarm

$\Downarrow$

k	1	2	3	4
$\Lambda(k)$	0	$1/3$	$3/2$	6

Under ML decisions

$$\Lambda(X) \begin{cases} > 1 & \text{declare } H_1 \text{ is true} \\ < 1 & \text{declare } H_0 \text{ is true.} \end{cases}$$

$$P_{\text{falsealarm}} = P(H_1 \text{ is declared} \mid H_0 \text{ is true})$$
$$= 0.2 + 0.1 = 0.3$$

You declared that there is an enemy aircraft but there is no enemy aircraft.

$$P_{\text{miss}} = P(H_0 \text{ is declared} \mid H_1 \text{ is true})$$
$$= 0.0 + 0.1 = 0.1$$

You declared that there is no enemy aircraft but there is one TT.

B. Maximum A posteriori Probability (MAP) rule.

You are now given "a posteriori probability" $P(H_1)$ and $P(H_0)$

With this additional information, you can detect the enemy aircraft better! Q: How?

EX) $P(H_1) = 0.2$ & $P(H_0) = 0.8$

Likelihood Matrix

	X=1	X=2	X=3	X=4
H_1	0.0	0.1	0.3	0.6
H_0	0.4	0.3	0.2	0.1

ML Decisions

Sum over each row = 1

$\Rightarrow$ Joint Probability Matrix

	X=1	X=2	X=3	X=4
H_1	0.0	0.02	0.06	0.12
H_0	0.32	0.24	0.16	0.08

MAP decisions

Sum over the matrix = 1

Def: Under the MAP decision rule: Given $X=j$ observed, decide H_i to maximize $P(H_i \mid X=j)$

Recall that it was $P(X=j \mid H_i)$ under the ML rule.

$$P(H_i \mid X=j) = \frac{P(\{X=j\} \cap H_i)}{P(X=j)}$$

$\Rightarrow$ Maximizing $P(H_i \mid X=j) \Leftrightarrow$ Maximizing $P(\{X=j\} \cap H_i)$

MAP decision with $\Lambda(k) = \dfrac{P(X=k \mid H_1)}{P(X=k \mid H_0)}$

$$\Lambda(X) \begin{cases} > \dfrac{P(H_0)}{P(H_1)} \Rightarrow \text{Declare } H_1 \text{ is true} \\ < \dfrac{P(H_0)}{P(H_1)} \Rightarrow \text{Declare } H_0 \text{ is true.} \end{cases}$$

Q: under what condition, ML decision = MAP decision?

Joint Probability Matrix

	X=1	X=2	X=3	X=4
H_1	0.0	0.02	0.06	0.12
H_0	0.32	0.24	0.16	0.08

MAP decisions

$$P_{\text{false alarm}} = P(H_1 \text{ declared} \mid H_0) = \frac{P(H_1 \text{ declared}, H_0)}{P(H_0)} = \frac{P(\{X=4\}, H_0)}{P(H_0)}$$

$$= 0.08 / P(H_0) = 0.1$$

$$P_{\text{miss}} = P(H_0 \text{ declared} \mid H_1) = \frac{P(H_0 \text{ is declared}, H_1)}{P(H_1)} = \frac{P(X \in \{0,1,2\}, H_1)}{P(H_1)}$$

$$= \frac{0.08}{P(H_1)} = 0.4$$

Prob that MAP detection is incorrect!

$$P_{\text{error}} = P(H_0) \times P_{\text{false alarm}} + P(H_1) P_{\text{miss}}$$

$$= 0.16$$

Ex) Two biased coins

1st coin ⇒ head comes out w.p. 0.6
2nd coin ⇒ head comes out w.p. 0.4

Experiment:

1. Pick a coin randomly w/o knowing which coin it is.
2. Toss the coin 1000 times.
3. Let X be the # of times that heads come out.

With the information on X, estimate whether the coin is the first or the second!

H_1 : The first coin is picked.
H_0 : The second coin is picked.

Conditioned on H_1, $X \sim$ Binomial $(1000, 0.6)$
Conditioned on H_2, $X \sim$ Binomial $(1000, 0.4)$

Find ML decision rule.

$$\Lambda(k) = \frac{P(X=k \mid H_1)}{P(X=k \mid H_0)} = \frac{\binom{1000}{k} 0.6^{k} 0.4^{1000-k}}{\binom{1000}{k} 0.4^{k} 0.6^{1000-k}} = \left(\frac{3}{2}\right)^{k} \left(\frac{2}{3}\right)^{1000-k}$$

$$= \left(\frac{3}{2}\right)^{2k-1000} \begin{cases} > 1 & \text{if } k > 500 \\ < 1 & \text{if } k < 500 \end{cases}$$

⇒ Declare H_1 is true if $X > 500$
H_0 is true if $X < 500$
anything if $X = 500$

MAP? Since $P(H_1) = P(H_0) = 0.5$, MAP = ML.

What if $P(H_1) = 2P(H_0)$?

$$\Lambda(k) \underset{H_0}{\overset{H_1}{\gtrless}} \frac{P(H_0)}{P(H_1)} = \frac{1}{2}$$

$$\Leftrightarrow \left(\frac{3}{2}\right)^{2k-1000} \gtrless \frac{1}{2}$$

$$\Leftrightarrow (2k-1000)\ln\left(\frac{3}{2}\right) \gtrless \ln\left(\frac{1}{2}\right)$$

$$\Longleftrightarrow k \gtrless \frac{\ln\left(\frac{1}{2}\right)}{2\ln\left(\frac{3}{2}\right)} + 500$$

MAP decision rule

Declare H_1 is true if $X > \frac{\ln\left(\frac{1}{2}\right)}{2\ln\left(\frac{3}{2}\right)} + 500$

H_0 is true if $X < \frac{\ln\left(\frac{1}{2}\right)}{2\ln\left(\frac{3}{2}\right)} + 500$

anything if $X = \frac{\ln\left(\frac{1}{2}\right)}{2\ln\left(\frac{3}{2}\right)} + 500$.