Lecture 4

Review :

Least square estimation

X — response

$Y_1, Y_2, \cdots Y_m \longrightarrow$ predictors

Criterion $E[X-\hat{X}]^2 \longrightarrow$ mean square error

Minimise $\hat{X}$ function of $(Y_1, \cdots, Y_m)$

$$\hat{X} = a_1 Y_1 + a_2 Y_2 + \cdots + a_m Y_m$$

Simplified version: $EX=0$
$E[Y_i]=0$, $i=1,2,\cdots,m$

$$E[X-\hat{X}]^2 = E[X-(a_1Y_1+a_2Y_2+\cdots+a_mY_m]^2$$
$$= Var(X) - 2\vec{a}[Cov(X,\vec{Y}) + \vec{a}^T \underbrace{Cov(\vec{Y})}_{\text{Note: "Covariance Matrix"}}\vec{a} = F(\vec{a})$$

$$\vec{a} = \begin{pmatrix} a_1 \\ a_2 \\ \vdots \\ a_m \end{pmatrix} \qquad \vec{Y} = \begin{pmatrix} Y_1 \\ \vdots \\ Y_m \end{pmatrix}$$

$$Cov(\vec{Y}) = \begin{pmatrix} Cov(Y_1,Y_1) & Cov(Y_1,Y_2)\cdots & Cov(Y_1,Y_m) \\ Cov(Y_2,Y_1) & Cov(Y_2,Y_2)\cdots & Cov(Y_2,Y_m) \\ \vdots & & \\ Cov(Y_m,Y_1) & \cdots & Cov(Y_m,Y_m) \end{pmatrix}$$

Target: Find such an "a"

Set $\dfrac{\partial F(\vec{a})}{\partial a_i} = 0 \quad i=1,2,\cdots,m$

$\Rightarrow \underbrace{Cov(\vec{Y})}_{\text{denoted as } V_{YY}}\ \vec{a} = \underbrace{Cov(X,\vec{Y})}_{\text{denoted as } V_{XY}}$

$$\vec{a} = \frac{V_{XY}}{V_{YY}}$$

(same as STA302)

CHAPTER 3
PROBABILITY

For any event $A \subset \Omega$. $P(A) = E[I(A)]$ $\qquad I(A) = \begin{cases} 1 & \text{if } A \in \Omega \\ 0 & \text{if } A \notin \Omega \end{cases}$

A little bit intro to Set Theory.

$A \subset \Omega$, $\bar{A} \longrightarrow$ complement of A.
$\bar{A} = \{\omega, \omega \notin A\}$

1. $I(\bar{A}) = 1 - I(A)$
2. If A & B are disjoint then $I(A \cup B) = I(A) + I(B)$
3. If $A_1, A_2, \cdots, A_n$ are disjoint, then $I(\bigcup_{i=1}^{n} A_i) = \sum_{i=1}^{n} I(A_i) = \max_{1 \le i \le n} I(A_i)$

4. If $A \subseteq B$, then $I(A) \le I(B)$

5. $I(\bigcap_{i=1}^{m} A_i) = \min_{1 \le i \le m} I(A_i)$

Properties of Probabilities

① $P(\Omega) = 1$
Proof: $I(\Omega) = 1 \xrightarrow{\text{Axiom 4}} E(I(\Omega)) = 1 \Rightarrow P(\Omega) = 1$

② $P(A) \ge 0$
because $I(A) \ge 0 \xrightarrow{\text{Axiom 1}} E(I(A)) \ge 0 \Rightarrow P(A) \ge 0$

③ If A & B are disjoint, then $P(A \cup B) = P(A) + P(B)$
Proof: $I(A \cup B) = I(A) + I(B)$

$\xrightarrow{\text{Axiom 3}} E(I(A \cup B)) = E(I(A)) + E(I(B))$

$P(A \cup B) = P(A) + P(B)$

④ If $A_1, A_2, \cdots, A_n, \cdots$ is a sequence of monotonic events
Def: A_i's are monotonic if
$A_1 \subset A_2 \subset \cdots \subset A_n \subset \cdots$ ①
$A_1 \supset A_2 \supset \cdots \supset A_n \supset \cdots$ ②
And in case ① $A_\infty = \bigcup_{i=1}^{\infty} A_i$
② $A_\infty = \bigcap_{i=1}^{\infty} A_i$

If we have a monotonic sequence of events. Then $P(A_\infty) = \lim_{i \to \infty} P(A_i)$

Proof: we only focus on ① since ② follows by the same argument.

Define r.v. $X_i = I(A_i)$

Since $A_i \subset A_{i+1}$

Hence $I(A_i) \le I(A_{i+1})$

$\Rightarrow X_i$ is a non-decreasing seq.

Axiom 5 implies $\lim_{i\to\infty} EX_i = E(\lim_{i\to\infty} X_i)$ (*)

LHS of (*) $= \lim_{i\to\infty} P(A_i)$

I'll show that $\lim_{i\to\infty} I(A_i) = I(A_\infty) = I(\bigcup_{i=1}^{\infty} A_i)$ (**)

a) If LHS of (**) $=0$

$\Rightarrow I(A_i)=0$ for all $i \Rightarrow \omega \notin A_i$, $i=1,2,\cdots$ $\omega \notin \bigcup_{i=1}^{\infty} A_i$

$\Rightarrow I(A_\infty)=0$

b). If LHS $=1$

$I(A_i)=1$ for some i

$\Rightarrow \omega \in$ some A_i

$\Rightarrow \omega \in \bigcup_{i=1}^{\infty} A_i \Rightarrow I(A_\infty)=1$

By (*) & (**) ④ holds.

Corollaries:

① If A & B are two events not necessarily disjoint. Then $P(A\cup B) = P(A) + P(B) - P(AB)$, $AB = A\cap B$

Note that $I(A\cup B) = I(A) + I(B) - I(AB)$

Why? (HW)

Hints: take expectations

$EI(A\cup B) = EI(A) + EI(B) - EI(AB)$

$\Rightarrow P(A\cup B) = P(A) + P(B) - P(A\cap B)$

② Inclusion-exclusion formula

Sps $A_1, A_2, \cdots, A_m$ are events

Then $P(\bigcup_{i=1} A_i) = \sum P(A_i) - \sum P(A_iA_j) + \sum P(A_iA_jA_k) - \cdots + (-1)^r \sum P(A_i \cdots A_r) + (-1)^m \sum P(A_i \cdots A_t)$

Proof by induction:

If $m=2$. proved.

$m=k$ ② holds

$m=k+1$, $P(\bigcup_{i=1}^{k+1} A_i) = P(B\cup A_{k+1}) = \cdots$ (straightforward)

③ If $A_1, \cdots, A_n, \cdots$ are disjoint events then

$P(\bigcup_{i=1}^{\infty} A_i) = \sum_{i=1}^{\infty} P(A_i)$

Proof: Let $B_i = \bigcup_{j=1}^{i} A_j$ Then B_i is a non-decreasing sequence.

Therefore, $P(\lim_{i\to\infty} B_i) = \lim_{i\to\infty} P(B_i)$

Note that $\lim B_i = \bigcup_{i=1}^{\infty} A_i$

$$\Rightarrow P\left(\bigcup_{i=1}^{\infty} A_i\right) = \lim_{i\to\infty} P\left(\bigcup_{j=1}^{i} A_j\right)$$

Ex: Sps we draw a card from a deck of 52 cards. Calculate the probability of drawing a diamond or an ace.

Let A = draw a diamond.

B = draw an ace.

Want $P(A\cup B)$

$P(A) = 1/4$

$P(B) = 1/13$

$P(A\cap B) = 1/52$

$P(A\cup B) = 1/4 + 1/13 - 1/52 = 4/13$

Ex:

Sps 4 letters are to be sent to 4 different people. The secretary who makes the envelopes carelessly, put the letters in randomly. What's the prob that at least one will get his/her letter?

Solution: . Inclu/exclu formula

Let A_i = the ith person gets his/her letter

$i = 1, 2, 3, 4$

Want $P(\cup A_i) = \sum P(A_i) - [P(A_1A_2) + P(A_2A_3) + P(A_3A_1) + P(A_1A_4) + P(A_2A_4) + P(A_3A_4)]$

$+ [P(A_1A_2A_3) + P(A_2A_3A_4) + P(A_1A_2A_4)] - P(A_1A_2A_3A_4)$

Take a look at $P(A_i) = 1/4$

$P(A_iA_j) = 1/12$

$P(A_iA_jA_k) = 1/24$

So $P(\cup A_i) = 1 - 1/12 \times 6 + 1/12 \times 3 - 1/24$

$= 1 - 1/2 + 1/4 - 1/24$

$= 3/4 - 1/24$

$= 17/24$

CHAPTER 4

SOME BASIC MODELS

A spatial model: Sps that there are N moleculars in a spatial region. The space is divided into M equal sized regions (cells).

Z_i = the cell the ith molecular occupies, $i = 1, 2, \cdots, N$.

Let $X = X(Z_1, Z_2, \cdots, Z_N)$

Then $EX = ?$

$$\boxed{EX = M^{-N} \sum_{i_1=1}^{M} \sum_{i_2=1}^{M} \cdots \sum_{i_N=1}^{M} X(i_1, i_2, \cdots, i_N)} \qquad (*)$$

Thm: If $(*)$ holds for every r.v. X. Then $Z_1, \cdots, Z_i$ must be uniformly distributed and

$$E\left[\prod_k H_k(Z_k)\right] = \prod_k E[H_k(Z_k)] \qquad \text{for any functions } H_k.$$

$(**)$

Proof: $i = 1, \cdots, N$
$j = 1, \cdots, M$

Let $X = I\{Z_i = j\}$
Take ? . Plug this X into $(*)$
We find that $P(Z_i = j) = 1/M$

Let $X = \prod_k^N H_k(Z_k)$

According to $(*)$ $EX = \frac{1}{M^N} \sum \cdots \sum X$

$$= \left[\frac{1}{M} \sum H_1\right]\left[\frac{1}{M} \sum H_2\right] \cdots \left[\frac{1}{M} \sum H_N\right]$$

$$= E[H_1(Z_1)]\, E[H_2(Z_2)] \cdots E[H_N(Z_N)]$$

∎

Let $X = X(Z_1, Z_2 \cdots)$

Def: For r.v. $X_1, \cdots, X_r$
we say $X_1, \cdots, X_r$ are independent if

$$E\left[\prod_r H_i(X_i)\right] = \prod_r E(H_i(X_i))$$

for any functions H_i.

Or equivalently, $X_1, \cdots, X_r$ r.v.'s

$$P(X_1 \in A_1, X_2 \in A_2, \cdots, X_r \in A_r) = P(X_1 \in A_1) P(X_2 \in A_2) \cdots P(X_r \in A_r)$$

for any $A_1, \cdots, A_r$