Mathematical Statistics

$1 \times 3 = 42$

1.4 Conditional Expectations (7.11.2019)

Definition of cpdf in elementary probability: $P(B|A) = \frac{P(A \cap B)}{P(A)}$, $P($

Question: what happens if $P(A) = 0$

$A = \{Y = c\}$; $c \in R$ and Y is random variable

In this section, the general definition of

- Conditional probability
- Expectation, and
- distribution

are introduced.

1.4.1 Conditional expectations:

Definition: X is integrable random variable on (Ω, F, P)

(i) $\mathcal{A}$ a sub-σ-field of $\mathcal{F}$ $\rightarrow$ $E(X|\mathcal{A})$ as unique as random variable

(a) $E(X|\mathcal{A})$ is measurable from $(\Omega, \mathcal{A})$ to $(R, \mathcal{B})$

(b) $\int_A E(X|\mathcal{A})dP = \int_A X dp$ for any $A \in \mathcal{A}$

(ii) $B \in \mathcal{F} \rightarrow P(B|\mathcal{A}) = E(I_B|\mathcal{A})$

(iii) Y is measurable from $(\Omega, \mathcal{F}, P)$ to $(\Lambda, \mathcal{G}) \rightarrow E(X|Y) = E[X|\sigma(Y)]$

Lemma 1.2: $(\Omega, \mathcal{F}) \xrightarrow[\text{is measurable}]{Y} (\Lambda, \mathcal{G})$

Z is function from $(\Omega, \mathcal{F})$ to R^k

$(\Omega, \sigma(Y)) \xrightarrow[Z \text{ is measurable}]{} (R^k, \mathcal{B}^k)$ if and only if $Z = h \circ Y$ measurable function

$\boxed{\text{Example}}$ X is an integrable random variable on $(\Omega, \mathcal{F}, P)$

$A_1, A_2, \ldots$ disjoints events on $(\Omega, \mathcal{F}, P)$ — $\cup A_i = \Omega$, $P(A_i) > 0$

$a_1, a_2, \ldots$ distinct real numbers

$Y = a_1 I_{A_1} + a_2 I_{A_2} + \ldots$

$\rightarrow E(X|Y) = \sum_{i=1}^{\infty} \frac{\int_{A_i} X dp}{P(A_i)} I_{A_i}$ we need to verify (a) and (b) in above definition

$A = \sigma(Y)$

$\sigma(Y) = \sigma(\{A_1, A_2, \dots\})$ then $\sum_{i=1}^{\infty} \frac{\int_{A_i} X\,dp}{P(A_i)} I_{A_i}$ is measurable on $(\Omega, \sigma(Y))$

$B \in \mathcal{B}$, $Y^{-1}(B) = \bigcup_{i: a_i \in B} A_i$

$$\int_{Y^{-1}(B)} X\,dp = \sum_{i: a_i \in B} \int_{A_i} X\,dp$$

$$= \sum_{i=1}^{\infty} \frac{\int_{A_i} X\,dp}{P(A_i)} P(A_i \cap Y^{-1}(B))$$

$$= \int_{Y^{-1}(B)} \left[\sum_{i=1}^{\infty} \frac{\int_{A_i} X\,dp}{P(A_i)} (I_{A_i}) \right] dp$$

h is Borel function on R and satisfying

$h(a_i) = \int_{A_i} X \frac{dp}{P(A_i)}$ $\rightarrow$ $E(X|Y) = h \circ Y$, $E(X|Y=y) = h(y)$

$A \in \mathcal{F}$, $X = I_A$ then

$$P(A|Y) = E(X|Y) = \sum_{i=1}^{\infty} \frac{P(A \cap A_i)}{P(A_i)} I_{A_i}$$

$P(A \cap A_i)/P(A_i) = P(A|A_i)$ if $\omega \in A_i$

---

proposition 1.9 – X: random n-vector

Y: random m-vector

(X, Y) has a joint pdf $f(x,y)$, $\nu \times \lambda$

↳ σ-finite measures on $(R^n, \mathcal{B}^n)$, $(R^m, \mathcal{B}^m)$

$g(x,y)$ is a Borel function on R^{n+m} for $E|g(X,Y)| < \infty$

then:

$$E[g(X,Y)|Y] = \frac{\int g(x,Y) f(x,Y)\,d\nu(x)}{\int f(x,Y)\,d\nu(x)}$$

$= h(Y) \rightarrow$ it is Borel by Fubini's theorem

$\boxed{\text{Proof}}$

Based on lemma 1.2 $h(Y)$ is Borel on $(\Omega, \sigma(Y))$ and because of Fubini's theorem

$$f_Y(y) = \int f(x,y)\,d\nu(x)$$

$$\int_{Y^{-1}(B)} h(Y)\,dP = \int_B h(y)\,dP_Y =$$

$$\int \frac{\int g(x,y) f(x,y)\,d\nu(x)}{\int f(x,y)\,d\nu(x)} f_Y(y)\,d\lambda(y)$$

$$= \int_{\mathbb{R}^n \times B} g(x,y)\, f(x,y)\, d\nu \times \lambda = \int_{\mathbb{R}^n \times B} g(x,y)\, dP(X,Y)$$

$$= \int_{Y^{-1}(B)} g(X,Y)\, dP$$

(X,Y) a random vector with joint pdf $f(x,y)$ the conditional pdf of X when $Y=y$

$$f_{X|Y}(x|y) = \frac{f(x,y)}{f_Y(y)} \quad : \quad f_Y(y) = \int f(x,y)\, d\nu(x)$$

proposition $\underline{1.10}$: $X, Y, X_1, X_2, \dots$ integrable random variables on $(\Omega, \mathcal{F}, P)$, $\mathcal{A}$ a sub-σ-field of $\mathcal{F}$

(i) if $X = c$, $c \in \mathbb{R} \rightarrow E(X|\mathcal{A}) = c$

(ii) if $X \leq Y \rightarrow E(X|\mathcal{A}) \leq E(Y|\mathcal{A})$

(iii) if $a \in \mathbb{R}$, $b \in \mathbb{R} \longrightarrow E(aX + bY|\mathcal{A}) = aE(X|\mathcal{A}) + bE(Y|\mathcal{A})$

(iv) $E[E(X|\mathcal{A})] = EX$

⋮

look the rest in page $\underline{40}$

$\boxed{\text{Example 1.22}}$ X a random variable on $(\Omega, \mathcal{F}, P)$, $EX^2 < \infty$

Y a measurable function from $(\Omega, \mathcal{F}, P)$ to $(\Lambda, \mathcal{G})$

$g(Y)$ is a predictor, i.e. $g \in \aleph = \{$all Borel functions g with $E[g(Y)]^2 < \infty\}$

predictor error : $E[X - g(Y)]^2$

$\Rightarrow E(X|Y)$ is the best predictor of X if

$$E[X - E(X|Y)]^2 = \min_{g \in \aleph} E[X - g(Y)]^2$$

$$E[X - g(Y)]^2 = E[X - E(X|Y) + E(X|Y) - g(Y)]^2$$

$$= E[X - E(X|Y)]^2 + E[E(X|Y) - g(Y)]^2 + 2E[X - E(X|Y)]E[E(X|Y) - g(Y)]$$

$$= \dots\dots \geq E[X - E(X|Y)]^2$$

1.4.2 independence :

Definition 1.7 : $(\Omega, \mathcal{F}, P)$ is a probability space

(i) $\mathcal{C}$ a collection of subset in $\mathcal{F}$. Events in $\mathcal{C}$ are independent if and only if for any positive integer n a distinct events

$A_1, \ldots, A_n$ in $\mathcal{C}$ $\rightarrow$

$$P(A_1 \cap A_2 \cap \ldots \cap A_n) = P(A_1) P(A_2) \ldots P(A_n)$$

(ii) collections $\mathcal{C}_i \subset \mathcal{F}$, $i \in I$ to be independent if and only if $\{A_i \in \mathcal{C}_i : i \in I\}$ are independent.

(iii) Random elements X_i, $i \in I$ are independent if and only if $\sigma(X_i)$, $i \in I$ are independent.

Lemma 1.3 : $\mathcal{C}_i$, $i \in I$ be independent collections of events. Each $\mathcal{C}_i$ has the property that if $A \in \mathcal{C}_i$ and $B \in \mathcal{C}_i$ then $A \cap B \in \mathcal{C}_i$. Then $\sigma(\mathcal{C}_i)$, $i \in I$ are independent.

For two events A, B with $P(A) > 0$, A and B are independent if $P(B|A) = P(B)$

$\boxed{\text{proposition } 1.11}$: X a random variable, $E|X| < \infty$

Y_i random k_i-vector, $i = 1, 2$

(X, Y_1) and Y_2 are independent then

$$E[X | (Y_1, Y_2)] = E(X | Y_1)$$