KTH, School of Electrical Engineering and Computer Science

Condition and decomposition of measures

Yuchao Li

January 24, 2019

Question 1

Q: Given $(\Omega, \mathscr{A}, \mathbf{P})$ and a random variable X, let $\mathbf{P}(E|X) = \mathbf{P}(E|\sigma(X))$ for $E \in \mathscr{A}$. Establish that there is a Borel measurable function $g : \mathbb{R} \to \mathbb{R}^+$ such that $\mathbf{P}(E|X) = g(X)$, $\mathbf{P}$-a.e., that is, as a function of $\omega \in \Omega$, $\mathbf{P}(E|X)$ has the form $\omega \to X(\omega) \to g\big(X(\omega)\big)$ ($\mathbf{P}$-a.e.). For $x \in \mathbb{R}$ and $E \in \mathscr{A}$, let $\mathbf{P}(E|X = x) = g(x)$; the conditional probability of E given $X = x$. Then, for any $E \in \mathscr{A}$ and $B \in \mathscr{B}$, prove that

$$\mathbf{P}(\{X \in B\} \cap E) = \int_B \mathbf{P}(E|X = x)\, d\mu_X.$$

A: Given a r.v. X, then r.v. X has distribution μ_X defined on $(\mathbb{R}, \mathscr{B})$. Namely for any $B \in \mathscr{B}$, we have $\mu_X(B) = \mathbf{P}(\{X \in B\})$. In addition, given $E \in \mathscr{A}$, we define a set function ν_E on $(\mathbb{R}, \mathscr{B})$ such that $\nu_E(B) = \mathbf{P}(\{X \in B\} \cap E)$. It can be shown that ν_E is in fact a measure on $(\mathbb{R}, \mathscr{B})$. Then since $\{X \in B\} \cap E \subseteq \{X \in B\}$, we have $\nu_E \leq \mu_X$ holds $\forall B \in \mathscr{B}$ and $\nu_E \ll \mu_X$. Besides, both ν_E and μ_X are finite, then by Radon-Nikodym theorem, we have $g = \frac{d\nu_E}{d\mu_X}$ μ_X a.e.. Namely, if $\hat{g} = \frac{d\nu_E}{d\mu_X}$, then $\hat{g} \neq g$ in a μ_X null set. Denote such set as $A \in \mathscr{B}$, then $\mu_X(A) = 0$, namely $\mu_X(A) = \mathbf{P}(X \in A) = 0$. Therefore, $\hat{g}(X)$ and $g(X)$, defined on Ω, can differ in $\{X \in A\}$, which is $\mathbf{P}$ null set.

Question 2

Q: Assume that X and Y are discrete random variables with joint pmf $p(x, y)$. Define

$$p(y|x) = \begin{cases} p(x,y)/p(x), & p(x) > 0, \\ 0, & \text{o.w..} \end{cases}$$

Show that

$$\mathbf{P}(\{Y = y\}|X = x) = p(y|x),$$

where $\mathbf{P}(\{Y \in B\}|X = x)$ is defined as in Question 1. Also determine $\mathbf{P}(\{Y = y\}|X)$.
A: Since both X and Y are discrete r.v., then there exists a countable set K such that $\mathbf{P}(X \in K) = 1$, then we can directly see that μ_X is a discrete measure. Then by definition, $\forall B \in \mathscr{B}$ we have

$$\mathbf{P}(\{Y = y\} \cap \{X \in B\}) = \sum_{x_i \in B\cap K} p(x_i, y). \tag{0.1}$$

On the other hand, by Question 5, we have that

$$\begin{aligned}\mathbf{P}(\{Y = y\} \cap \{X \in B\}) &= \int_B \mathbf{P}(\{Y = y\}|X = x)d\mu_X \\ &= \sum_{x_i \in B\cap K} \mathbf{P}(\{Y = y\}|X = x_i)p(x_i).\end{aligned} \tag{0.2}$$

Comparing Eqs. (0.1) and (0.2), we can see that $\mathbf{P}(\{Y = y\}|X = x) = p(y|x)$.

QUESTION 3

Q: Assume that X and Y are jointly absolutely continuous random variables with joint pdf $f(x, y)$. Define

$$f(y|x) = \begin{cases} f(x, y)/f(x), & f(x) > 0, \\ 0, & \text{o.w..}\end{cases}$$

Show that for each $B \in \mathscr{B}$,

$$\mathbf{P}(\{Y \in B\}|X = x) = \int_B f(y|x)dy,$$

where $\mathbf{P}(\{Y \in B\}|X = x)$ is defined as in Question 1. Also determine $\mathbf{P}(\{Y \in B\}|X)$.
A: Since both X and Y are continuous r.v.'s, then we can see that μ_X and μ_Y are absolutely measure. Then by definition, $\forall A, B \in \mathscr{B}$ we have

$$\mathbf{P}(\{Y \in B\} \cap \{X \in A\}) = \int_A \int_B f(x, y)dydx. \tag{0.3}$$

On the other hand, by Question 5, we have that

$$\mathbf{P}(\{Y \in B\} \cap \{X \in A\}) = \int_A \mathbf{P}(\{Y \in B\}|X = x)d\mu_X. \tag{0.4}$$

Denote $\nu(A) = \mathbf{P}(\{Y \in B\} \cap \{X \in A\})$ as a set function defined on $(\mathbb{R}, \mathscr{B})$, it can be seen that ν in fact is a measure and $\nu \ll \mu_X$. Besides, we have $\frac{d\nu}{d\mu_X}(x) = \mathbf{P}(\{Y \in B\}|X = x)$ μ_X a.e.. Due to properties of Radon-Nikodym derivative, we have $\nu \ll \mu_X \ll \lambda$, and $\frac{d\nu}{d\lambda} = \frac{d\nu}{d\mu_X}\frac{d\mu_X}{d\lambda}$, refer to [Note 12], then we have

$$\mathbf{P}(\{Y \in B\} \cap \{X \in A\}) = \int_A \mathbf{P}(\{Y \in B\}|X = x)f(x)dx \tag{0.5}$$

Alternatively, we can also apply Proposition 9.1, P. 318, [1], which in fact can be proved by the above arguments. Comparing Eqs. (0.3) and (0.5), the proof is done.

QUESTION 4

Q: Given $(\Omega, \mathscr{A}, \mathbf{P})$ and a r.v. X such that $E(X^2) < \infty$. Given $\mathscr{G} \subset \mathscr{A}$ let $\mathscr{F} = \{\text{all } \mathscr{G}\text{-measurable functions}\}$. Show that $\hat{Y}(\omega) = E[X|\mathscr{G}]$ is a $\mathbf{P}$-a.e. unique solution to the problem: find $Y \in \mathscr{F}$ to minimize $E[(X-Y)^2]$.
A: Denote $\hat{Y} = E[x|\mathscr{G}]$ and $\tilde{Y} = Y - \hat{Y}$ where $Y \in \mathscr{F}$ is some arbitrary r.v.. Then we have

$$\begin{aligned}
E[(X-Y)^2] &= E[E[(X-Y)^2|\mathscr{G}]] \\
&= E[E[X^2 - 2XY + Y^2|\mathscr{G}]] \\
&= E[E[X^2|\mathscr{G}]] + E[E[Y^2|\mathscr{G}]] - 2E[E[XY|\mathscr{G}]] \\
&= E[X^2] + E[Y^2] - 2E[YE[X|\mathscr{G}]] \\
&= E[X^2] + E[Y^2] - 2E[Y\hat{Y}] \\
&= E[X^2] + E[(\hat{Y}+\tilde{Y})^2 - 2\hat{Y}(\hat{Y}+\tilde{Y})] \\
&= E[X^2] + E[\hat{Y}^2 + \tilde{Y}^2 + 2\hat{Y}\tilde{Y} - 2\hat{Y}^2 - 2\hat{Y}\tilde{Y}] \\
&= E[X^2] + E[\tilde{Y}^2 - \hat{Y}^2] \\
&= E[X^2] - E[\hat{Y}^2] + E[\tilde{Y}^2]
\end{aligned}$$

The minimal is obtained if $E[\tilde{Y}^2] = 0$.

REFERENCES

[1] John McDonald and Neil A. Weiss, *A course in real analysis*, 2nd Edition, 2012.

[2] Timo Koski, *Lecture notes: Probability and random processes at KTH*, 2017.