Feynman Liang
STAT 205A
Lecture #13
10/6/2016

1 Conditional Distributions

Definition 1.1. Given measurable spaces $(S_1, \mathcal{S}_1)$ and $(S_2, \mathcal{S}_2)$, define the *product measurable space*

$$(S_1 \times S_2, \mathcal{S}_1 \otimes \mathcal{S}_2) = \sigma(A \times B : A \in \mathcal{S}_1, B \in \mathcal{S}_2) \tag{1.1}$$

Random variables

$$X = (\Omega, \mathcal{F}, P \mapsto (S_1, \mathcal{S}_1)) \tag{1.2}$$
$$Y = (\Omega, \mathcal{F}, P \mapsto (S_2, \mathcal{S}_2)) \tag{1.3}$$

(X, Y) is a RV with values in $S_1 \times S_2$, has a distribution μ: a PM on $S_1 \times S_2$.
X has a distribution μ_1: a PM on S_1.
What is conditional distribution of Y given G?
Old write-up on web page: If $S_1 = S_2 = S$ countable, then $P(Y = y \mid X = x) = f(y \mid x)$ has properties:

- $f(y \mid x) \geq 0$
- $\sum_y f(y \mid x) = x$ for all x
- $P(X = x, Y = y) = P(X = x)P(Y = y \mid X = x)$

Definition 1.2. A *kernel* Q from S_1 to S_2 is a map $Q : (S_1 \times \mathcal{S}_2) \to [0, 1]$ such that

(a) For fixed s_1, $B \mapsto Q(s_1, B)$ is a PM on S_2

(b) For fixed $B \in \mathcal{S}_2$, $s_1 \mapsto Q(s_1, B)$ is a *measurable* function $S_1 \to \mathbb{R}$

Warning: for $h : S_1 \times S_2 \to \mathbb{R}$

(a) h is measurable

(b)

$$\forall s_1 : s_2 \to h(s_1, s_2) \text{ is measurable } S_2 \to \mathbb{R} \tag{1.4}$$
$$\forall s_2 : s_1 \to h(s_1, s_2) \text{ is measurable } S_1 \to \mathbb{R} \tag{1.5}$$

(a) $\implies$ (b) but not vice versa.

Example 1.3. $S_1 = S_2 = [0, 1], h(x, x) = 1_{x \in A}, h(x, y) = 0$. Non-measurable on $A \subset [0, 1]$

We interpret $P(Y \in B | X = s_1) = Q(s_1, B)$.

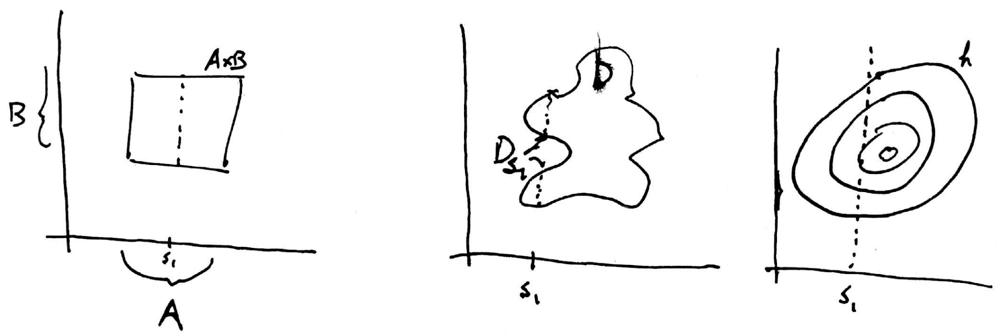
Proposition 1.4. *Given PM μ on $S_1 \times S_2$, a PM μ_1 on S_1, and a kernel Q from S_1 to S_2, TFAE:*

BR1 $\mu(A \times B) = \int_A Q(s_1, B)\mu_1(ds_1)\ \forall A \in \mathcal{S}_1, \forall B \in \mathcal{S}_2$

BR2 $\mu(D) = \int_{S_1} Q(s_1, D_{s_1})\mu_1(ds_1)\ \forall D \in \mathcal{S}_1 \otimes \mathcal{S}_2$

BR3 $\int_{S_1 \times S_2} h(s_1, s_2)\mu(d \underbrace{\tilde{s}}_{\tilde{s}=(s_1,s_2)}) = \int_{S_1} \left(\int_{S_2} h(s_1, s_2) Q(s_1, ds_2) \right) \mu_1(ds_1)$ *provided $h \geq 0$ or h is μ-integrable*

Here, $D_{s_1} := \{s_2 : (s_1, s_2) \in D\} \subset S_2$.



Lemma 1.5. *For each $D \in \mathcal{S}_1 \otimes \mathcal{S}_2$,*

(a) $D_{s_1} \in \mathcal{S}_2$ for all $s_1 \in \mathcal{S}_1$

(b) map $s_1 \mapsto Q(s_1, D_{s_1})$ is measurable

Proof. Let $\mathcal{D}$ be collection of all D satisfying (i,ii).

$\mathcal{D}$ is a λ-class. For $D \in \mathcal{D}$ have $(D^c)_{s_1} = \{s_2 : (s_1, s_2) \notin D\} = (D_{s_1})^c \in \mathcal{S}_2$ so $D^c \in \mathcal{D}$. Can also show $\mathcal{D}$ closed under increasing limits using closure of $\mathcal{S}_2$ and PM $Q(s, \cdot)$ under increasing limits:

$$D^n \uparrow D \implies D^n_{s_1} \uparrow D_{s_1} \implies Q(s_1, D^n_{s_1}) \uparrow Q(s_1, D_{s_1}) \tag{1.6}$$

Let $\mathcal{I} = \{A \times B : A \in S_1, B \in S_2\}$ be generated by rectangles. As $D = A \times B$ implies that $\forall s_1 \in S_1 : D_{s_1} = B \in \mathcal{S}_2$, we have $\mathcal{I} \subset \mathcal{D}$.

By definition 1.1, $\mathcal{S}_1 \otimes \mathcal{S}_2 = \sigma(\mathcal{I})$. By Dynkin's $\pi - \lambda$ theorem, $\sigma(\mathcal{I}) \subset \mathcal{D}$. $\square$

Theorem 1.6 (Easy theorem). *Given a PM μ_1 on S_1, given a kernel Q from S_1 to S_2, the definition*

$$\mu(D) = \int_{S_1} Q(s_1, D_{s_1})\mu_1(ds_1) \quad D \in \mathcal{S}_1 \otimes \mathcal{S}_2 \tag{1.7}$$

defines a PM μ on $S_1 \times S_2$.

Proof.

$$\mu(S_1 \times S_2) = \int_{S_1} \underbrace{\cancelto{1}{Q(s_1, S_2)}}_{\text{PM}} \mu_1(ds_1) = \mu_1(S_1) = 1 \tag{1.8}$$

$\mu(E) \geq 0$ and $\mu(\emptyset = 0)$ follow from $Q(s_1, \cdot)$ and μ_1 being measures, as does countable additivity: if $\{E_i\}$ are pairwise disjoint

$$\mu\left(\bigcup_i E_i\right) = \int_{S_1} Q(s_1, (\cup_i E_i)_{s_1})\mu_1(ds_1) = \int_{S_1} \sum_i Q(s_1, (E_i)_{s_1})\mu_1(ds_1) = \sum_i \mu(E_i) \tag{1.9}$$

□

Theorem 1.7 (Hard theorem). *Given PM μ on $S_1 \times S_2$, define marginal PM μ_1 on S_1 by $\mu_1(A) = \mu(A \times S_2)$.*

If S_2 is a Borel space, *then $\exists$ kernel Q from S_1 to S_2 such that proposition 1.4 hold.*

Proof. Fix $B \in \mathcal{S}_2$. Consider $\nu(A) := \mu(A \times B)$, $A \in \mathcal{S}_1$. ν is a (sub-probability) measure on $\mathcal{S}_1$.

$$\nu(A) = \mu(A \times B) \leq \mu(A \times \mathcal{S}_2) = \mu_1(A) \implies \nu \ll \mu_1 \tag{1.10}$$

Consider the Radon-Nikodyn density

$$\frac{d\nu}{d\mu_1}(s_1) = Q(s_1, B) \quad (\text{def of } Q(s_1, B)) \tag{1.11}$$

which satisfies the properties

$$s_1 \mapsto Q(s_1, B) \text{ is measurable} \tag{1.12}$$

$$\nu(A) = \int_A \frac{d\nu}{d\mu_1}(s_1)\mu_1(ds_1) \iff \mu(A \times B) = \int_{S_1} Q(s_1, B)\mu_1(ds_1) \quad \forall A \in \mathcal{S}_1 \tag{1.13}$$

These hold for any $B \in \mathcal{S}_2$, so $Q(s_1, B)$ satisfies the first property of a kernel $Q : S_1 \times \mathcal{S}_2 \to \mathbb{R}$ and BR1.

It remains to show that Q satisfies the second property of a kernel, namely $B \mapsto Q(s_1, B)$ is a PM on $\mathcal{S}_2$ $\forall s_1 \in S_1$.

Issue: If $h_1 \overset{\text{a.e.}}{=} h_2$ (wrt μ_1), then $\int_A h_1 d\mu_1 = \int_A h_2 d\mu_1$.

As S_2 is a Borel space, wlog assume $S_2 = \mathbb{R}$. For each rational $r \in \mathbb{Q}$ do construction for $B = (-\infty, r]$.

Write $F(s, r) = Q(s_1, (-\infty, r])$. Note

$$s_1 \mapsto F(s_1, r) \text{is measurable} \tag{1.14}$$

$$\mu(A \times (-\infty, r_1]) = \int_A F(s_1, r)\mu_1(ds_1) \quad \forall A \in \mathcal{S}_1 \tag{1.15}$$

Consider $r_1 < r_2 \in \mathbb{Q}$. For any $A \in \mathcal{S}_1$

$$\mu(A \times (r_1, r_2]) = \int_A (F(s_1, r_2) - F(s_1, r_1))\mu(ds_1) \geq 0 \tag{1.16}$$

so

$$F(s_1, r_2) \geq F(s_1, r_1) \quad \text{a.e.in } S_1 \tag{1.17}$$

Modify $F(s_1, r)$ **on null-set to make monotone on rationals**: Redefine $F(s_1, r)$ over the null set $\{s_1 : F(s_1, r_2) < F(s_1, r_1)\}$ such that $F(s_1, r) = \Phi(r)$ is monotone for $r \in \mathbb{Q}$. Repeat for all rational pairs (r_1, r_2) to get a version of $F(s_1, r)$ such that for any $s_1 \in A$, $r \mapsto F(s_1, r)$ is monotone on rationals. (A)

Since F is monotone on rationals and $F(s_1, r) = Q(s_1, (-\infty, r])$ where $B \mapsto Q(s, B)$ is a probability measure

$$\left.\begin{array}{l} \lim_{r\uparrow\infty} F(s_1, r) = 1 \ \forall s_1 \\ \lim_{r\downarrow\infty} F(s_1, r) = 0 \ \forall s_1 \end{array}\right\} (B) \tag{1.18}$$

Consider $r_n \downarrow r \in \mathbb{Q}$.

$$\left.\begin{array}{l} \mu(A \times (r_1, r_n]) \to 0 \ \forall A \\ F(s_1, r_n) \downarrow F(s_1, r) \text{ a.e.} \end{array}\right\} (C) \tag{1.19}$$

Modify on another null-set $r_n \downarrow r \in \mathbb{Q} \implies F(s_1, r_n) \downarrow F(s_1, r) \ \forall s_1$.

Remark 1.8 (Deterministic Fact). If $r \mapsto F(r)$ rational has properties (A), (B), (C), then

$$\hat{F}(x) = \lim_{\substack{r\downarrow x \\ r>x \\ r\in\mathbb{Q}}} F(r) \tag{1.20}$$

is a distribution function

$$\hat{F}(r) = F(r) \tag{1.21}$$

Use fact to define

$$\hat{F}(s_1 x, x) = \lim_{r\downarrow x} f(s_1, r) \ \forall x \in \mathbb{R} \tag{1.22}$$

where

$$s_1 \to \hat{F}(s_1, x) \text{ is measurable} \tag{1.23}$$

$$x \to \hat{F}(s_1, x) \text{ is a dist function} \tag{1.24}$$

Define Q by $Q(s_1, \cdot)$ is the PM with distribution function $F(s_1, x)$. $\square$