Feynman Liang
STAT 205A
Lecture #15
10/13/2016

1 More RVs and distributions

Corollary 1.1. *Given a PM μ on $S \times \mathbb{R}$.*

Given a RV $X : \Omega \to S$ where $dist(X) = \mu_1 =$ marginal of μ.

Given a RV $U : \Omega \to [0,1]$, $dist(U)$ is $U[0,1]$, U independent of X.

Then $\exists f : S \times [0,1] \to \mathbb{R}$ such that, with $Y = f(X,U)$, then $dist(X,Y) = \mu$.

Proof. Let Q be a kernel from S to $\mathbb{R}$ associated with μ.

Let $f(s,u)$ be the inverse distribution function of PM $Q(s,\cdot)$

$f(s,U)$ has distribution $Q(s,\cdot)$ since $Q(s,B) = \lambda\{u : f(s,u) \in B\}$ for all $B \in \mathcal{B}$ ($U \sim U[0,1]$ has distribution $\lambda =$ Lebesgue measure).

Check that this f works:

$$P(X \in A, Y \in B) = P(X \in A, f(X,U) \in B) \tag{1.1}$$
$$= \int\int 1_{X\in A} 1_{f(X,U)\in B}\mu(dx) \otimes \lambda(du) \tag{1.2}$$
$$\underset{\text{Fubini}}{=} \int 1_{X\in A}\left(\int 1_{f(X,U)\in B}\lambda(du)\right)\mu(dx) = \int_A Q(x,B)\mu(dx) \tag{1.3}$$
$$\underset{\text{Def of } Q}{=} \mu(A\times B) \tag{1.4}$$

□

This theorem, along with the following, justifies Markov chains in MT setting.

Notation: Let $1 \le m < n < \infty$, $\tilde{\pi}_{n,m} : (x_1, x_2, \cdots, x_n) \to (x_1, \cdots, x_m)$, $\pi_{n,m}$ the associated map $\mathcal{P}(\mathbb{R}^n) \to \mathcal{P}(\mathbb{R}^m)$ such that $\text{dist}(X_1,\cdots,X_n) \to \text{dist}(X_1,\cdots,X_m)$.

Theorem 1.2 (Kolmogorov Extension (Consistency) Theorem)**.** *Given PMs μ_n on $\mathbb{R}^n$, $1 \le n < \infty$, which are consistent in the sense $\pi_{n,m}\mu_n = \mu_m$, $1 \le m < n < \infty$.*

Then $\exists$ a PM μ_∞ on $\mathbb{R}^\infty$ such that $\pi_{\infty,m}\mu_\infty = \mu_m$, $1 \le m < \infty$.

Proof. Take $U_1, U_2, \cdots$ independent $U[0,1]$. Define $X_1 = F_{\mu_1}^{-1}(U_1)$. Inductively, suppose we have defined $X = (X_1, \cdots, X_n)$ as functions of $(U_1, \cdots, U_n)$ such that $\text{dist}(X_n) = \mu_n$.

We can show $\exists f_{n+1}$ such that defining $X_{n+1} = f_{n+1}(X_n, U_{n+1})$ we have $\text{dist}(\tilde{X_{n+1}} = (\tilde{X_n}, X_{n+1})) = \mu_{n+1}$. To do that, apply the corollary with $S = \mathbb{R}^n$, $X = \tilde{X}_n$, $U = U_{n+1}$, $\mu = \mu_{n+1}$ on $\mathbb{R}^n \times \mathbb{R}$.

This constructs infinite $(X_n, 1 \le n < \infty)$. Define $\mu_\infty = \text{dist}(X_n, 1 \le n < \infty)$. □

Example 1.3. Given measurable $h : \mathbb{R} \to \mathbb{R}$ and a PM μ that is invariant under h, i.e. $\text{dist}(X) = \mu \implies \text{dist}(h(X)) = \mu$.

For each n, take $\text{dist}(X_n) = \mu$. Define $X_i = h(X_{i+1}), 1 \le i \le n-1, \mu_n = \text{dist}(X_1, \cdots, X_n)$.

Theorem $\implies \mu_\infty = \text{dist}(Y_1, Y_2, \cdots)$ such that $\text{dist}(Y_1, \cdots, Y_n) = \text{dist}(X_1, \cdots, X_n) \quad \forall n$. $Y_i = h(Y_{i+1})$ for all $1 \le i < \infty$.

2 Intermission: example relevant to data

Hypothesis: Probabilities from gambling odds are indistinguishable from "true probabilities" as formalized in math.

Question: Does this hyp make predictions that can be checked against data?

Model: $Z_1 =$ point difference at half time. $Z_2 =$ point differencei n second half. Home team wins $\iff$ $Z_1 + Z_2 > 0$. Assume $Z_n \overset{d}{=} -Z_1$ (symmetric), $Z_1 \perp Z_2$, Z_1 has continuous dist.

$$\begin{aligned}
P(\text{home team wins} \mid Z_1 = z) &= P(Z_2 \geq -z \mid Z_1 = z) && (2.1)\\
&= P(Z_2 \geq -z) \text{ by indep.} && (2.2)\\
&= P(Z_2 \leq z) \text{ by symmetry} && (2.3)\\
&= F_Z(z) && (2.4)
\end{aligned}$$

$$P(\text{home team wins} \mid Z_1) = F_2(z_1) \overset{d}{=} U[0,1] \tag{2.5}$$

3 Conditional Expectation in measure theory setting

3.1 Undergraduate version

X, Y $\mathbb{R}$-valued, A an event.

$\mathbb{E}X$ is a number.

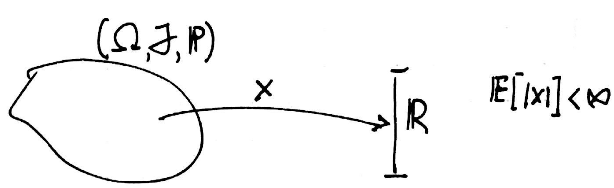
$\mathbb{E}[X \mid A]$ is a number.

$\mathbb{E}[X \mid Y = y]$ is a number depending on y i.e. a function of y, say $= h(y)$.

Write $E[X \mid Y] = h(Y)$, view as a RV. Useful because of tower property: $\mathbb{E}\mathbb{E}[X \mid Y] = \mathbb{E}X$.

Another way to define it is as a best least-squares estimator.

3.2 Measure theory setup



Consider a sub-σ-field $\mathcal{G} \subset \mathcal{F}$. We will define $\mathbb{E}[X \mid \mathcal{G}]$ to be a certain $\mathcal{G}$-measurable RV.

Suppose know information in $\mathcal{G}$. Fair stake now $= Y$, say

(a) Y is $\mathcal{G}$-measurable

(b) $\mathbb{E}Y1_G = \mathbb{E}X1_G \quad \forall G \in \mathcal{G}$

(b) is because of the following betting strategy: Choose $G \in \mathcal{G}$, bet if G happens, not if G^c happens, gain $(X - Y)1_G$. Fair $\iff$ $\mathbb{E}(\text{gain}) = 0\ \forall$strategies $\iff$ $\mathbb{E}(X - Y)1_G = 0\ \forall G$.

Define $\mathbb{E}[X \mid \mathcal{G}]$ to be *the* RV Y satisfying (a) and (b).

Existence: For $G \in \mathcal{G}$, define $\nu(G) = \mathbb{E}X1_G$. $P(G) = 0 \implies \nu(G) = 0$. $\nu \ll P$ as PMs on $(\Omega, \mathcal{G})$.

RN works for ν a signed-measure. Defining property of RN density is (b).

Radon-Nikodym $\implies \exists$ density $\frac{d\nu}{dP}(\omega) = Y(\omega)$.

Radon-Nikodym $\implies$ $\mathcal{G}$-measurable.

Uniqueness:

Lemma 3.1. *If Y is $\mathcal{G}$-measurable, if $\mathbb{E}|Y| < \infty$. If $\mathbb{E}(Y1_G) \geq 0, \forall G \in \mathcal{G}$, then $Y \geq 0$ a.s.*

Proof. If not, $G := \{Y < 0\}$ has $P(G) > 0$ and $\mathbb{E}Y1_G < 0$. Contradiction. □

Corollary 3.2. *If Y_1 and Y_2 each satisfy (a) and (b), then $Y_1 \overset{a.s.}{=} Y_2$.*

Proof.

$$\mathbb{E}(Y_1 - Y_2)1_G = 0\ \forall G \underset{\text{Lemma}}{\implies} Y_1 \geq Y_2 \text{ a.s. and } Y_1 \leq Y_2 \text{ a.s.} \implies Y_2 \overset{\text{a.s.}}{=} Y_2 \tag{3.1}$$

□

Lemma 3.3 (Technical Lemma). *(a) If $Z = \mathbb{E}[X \mid Y]$ then $\mathbb{E}[VZ] = \mathbb{E}[VX]$ for all bounded $\mathcal{G}$-measurable V.*

(b) If Z is $\mathcal{G}$-measurable, then to prove $Z = \mathbb{E}[X \mid Y]$ it suffices to prove $\mathbb{E}Z1_A = \mathbb{E}X1_A\ \forall A \in \mathcal{A}$ where $\mathcal{A}$ is a π-class such that $\mathcal{G} = \sigma(\mathcal{A})$.

Proof. (a) Use def for $V = 1_G$, monotone class theorem

(b) Dynkin $\pi - \lambda$ lemma

□