Feynman Liang
STAT 205A
Lecture #6
9/13/2016

1 Independence (undergrad)

Events A and B are *independent* if $P(A \cap B) = P(A)P(B)$.

R.V.s X, Y independent if $P(X \leq x, Y \leq y) = P(X \leq x)P(Y \leq y)$.

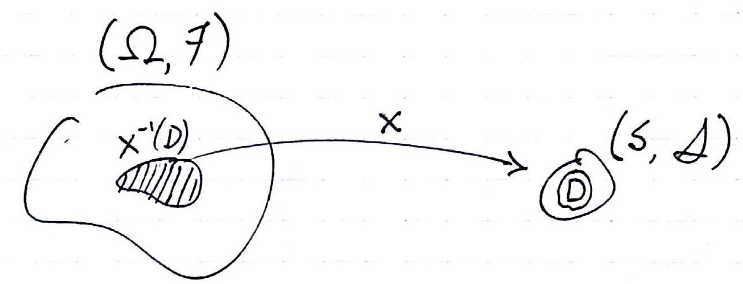
Idea: knowing the value of X doesn't change the probabilities for Y

2 Independence (MT version)

$(\Omega, \mathcal{F}, P)$ a probability space. Consider $\mathcal{B}_1, \mathcal{B}_2$ sub-σ-fields of $\mathcal{F}$.

Definition 2.1. Call σ-fields $\mathcal{B}_1$ and $\mathcal{B}_2$ *independent* if

$$P(B_1 \cap B_2) = P(B_1)P(B_2) \quad \forall B_i \in \mathcal{B}_i \tag{2.1}$$



X is measurable $\implies X^{-1}(D) \in \mathcal{F}$ for all $D \in \mathcal{S}$. The collection $\{X^{-1}(D) : D \in \mathcal{S}\}$ is a sub-σ-field of $\mathcal{F}$, call it $\sigma(X)$ "the σ-field generated by X."

Definition 2.2. Call RVs X_1 and X_2 independent if $\sigma(X_1)$ and $\sigma(X_2)$ are independent.

Theorem 2.3. *For RVs X_1, X_2 with X_i taking values in $(S_i, \mathcal{S}_i)$, the following are equivalent:*

(i) *X_1 and X_2 are independent*

(ii) $P(X_1 \in B_1, X_2 \in B_2) = P(X_1 \in B_1)P(X_2 \in B_2)$ *for all* $B_i \in \mathcal{S}_i$

(iii) $P(X_1 \in B_1, X_2 \in B_2) = P(X_1 \in B_1)P(X_2 \in B_2)$ *for all* $B_i \in \mathcal{A}_i$ *where* $\mathcal{A}_i$ *is a* π*-class,* $\sigma(\mathcal{A}_i) = \mathcal{S}_i$

(iv) $\mathbb{E}[h_1(X_1)h_2(X_2)] = (\mathbb{E}h_1(X_1))(\mathbb{E}h_2(X_2))$ *for all bounded measurable* $h_i : \S_i \in \mathbb{R}$

Comments:

- (iv) extends to integrable $h_i(X_i)$
- If X_i are $\mathbb{R}$-valued, (iii) can be used to show independence is equivalent to $P(X_1 \leq x_1, X_2 \leq x_2) = P(X_1 \leq x_1)P(X_2 \leq x_2)$ for all $x_i \in \mathbb{R}$
- The fact:

  > if X_1, X_2 are independent, then $g_1(X_1), g_2(X_2)$ independent (arbitrary measurable g_i)

  is true because $\sigma(g(X)) \leq \sigma(X)$.

Proof outline. (iv) $\implies$ (ii) $\implies$ (iii) special cases
(ii) $\implies$ (iv) by "monotone class argument":
(iv) holds for $h_i = 1_{B_i}$ indicator functions
$\therefore$ holds for h_i simple functions
$\therefore$ holds for h_i bounded measurable functions
Want to use Dynkin's $\pi - \lambda$ Lemma:

Step 1 Fix $B_2 \in \mathcal{A}_2$. Consider the collection

$$\mathcal{L} = \{A \in \mathcal{S}_1 : P(X_1 \in A, X_2 \in B_2) = P(X_1 \in A)P(X_2 \in B_2) \tag{2.2}$$

Check $\mathcal{L}$ is a λ-class.

By hypothesis, $\mathcal{L} \supset \mathcal{A}_1$. Dynkin's lemma implies $\mathcal{L} = \mathcal{S}_1$

Step 2 Consider

$$\mathcal{L}' = \{B_2 \in \mathcal{S}_2 : P(X_1 \in B_2, X_2 \in B_2) = P(X_1 \in B1, X_2 \in B_2) \quad \forall B_1 \in \mathcal{S}_1\} \tag{2.3}$$

Check $\mathcal{L}'$ is a λ-class (use linearity property).

By step 1, $\mathcal{L}' \supset \mathcal{A}_2$.

By Dynkin's, $\mathcal{L}' \supset \sigma(A_2) = \mathcal{S}_2 \implies$ (ii)

□

Definition 2.4. $\mathcal{B}_1, \mathcal{B}_2, \cdots, \mathcal{B}_n$ are *(mutually) independent* means

$$P(\cap_{i=1}^n B_i) = \prod_{i=1}^n P(B_i) \quad \forall B_i \in \mathcal{B}_i \tag{2.4}$$

This is *stronger* than pairwise independence.

Example 2.5. X, Y for die throws, events $\{X = 3\}$, $\{Y = 6\}$, $\{X = Y\}$. These events are only pairwise independent, not mutually independent.

Example 2.6. X_1, X_2 independent uniform on $\{0, 1, \cdots, n-1\}$. Define $X_3 = X_1 + X_2$ modulo n. Then (X_1, X_2, X_3) are pairwise independent, not mutually independent.

Claim. *If $X_1, \cdots, X_k, \cdots, X_n$ independent, then $f(X_1, \cdots, X_k)$ and $g(X_{k+1}, \cdots, X_n)$ are independent for arbitrary measurable functions f and g.*

Exercise 2.7. Formalize and verify. "Hereditary property of independence."

Exercise 2.8. To show that events $\{A_i\}_{i=1}^n$ are independent, suffices to show

$$P(\cap_{i\in\mathcal{I}} A_i) = \prod_{i\in\mathcal{I}} P(A_i) \quad \forall \mathcal{I} \subset \{1, 2, \cdots, n\} \tag{2.5}$$

3 Real-valued Random Variabls

Let X_i, Y_i be real-valued random variables.

Know that $X_n \to X$a.s. means $P(\{\omega : X_n(\omega) \to X(\omega)\text{a.s.}n \to \infty\}) = 1$.

Definition 3.1 (Convergence in probability). $X_n \to_p X$ means

$$\lim_{n\to\infty} P(|X_n - X| > \epsilon) = 0, \quad \forall \epsilon > 0 \tag{3.1}$$

Definition 3.2 (Convergence in L^p). For $l \leq p < \infty$, say $X_n \to X$ *in* L^p or $X_n \to^{L_p} X$ to mean

$$\mathbb{E}|X_n - X|^p = \|X_n - X\|_p \to 0 \text{ as } n \to \infty \tag{3.2}$$

(and $\mathbb{E}|X_n|^p < \infty$ for all n).

Lemma 3.3. *If $X_n \to X$ in L^p, then $X_n \to_p X$.*

Proof. Use general form of Markov's inequality $\phi(x) = |x|^p$ applied to $X_n - X$.

$$\implies P(\|X_n - X\| > \epsilon) \leq \frac{\mathbb{E}|X_n - X|^p}{\epsilon^p} \to 0 \text{ as } n \to \infty \tag{3.3}$$

□

Definition 3.4 (Variance). If $\mathbb{E}X^2 < \infty$, define the *variance*

$$\text{Var}(X) = \mathbb{E}X^2 - (\mathbb{E}X)^2 = \mathbb{E}(X - \mathbb{E}X)^2 \tag{3.4}$$

Proposition 3.5. *If $(X_i)_{i=1}^n$ independent, then $Var(\sum_i X_i) = \sum_i Var(X_i)$*

Definition 3.6. If $\mathbb{E}X_i^2 \leq \infty$, $\mathbb{E}X_1X_2 = (\mathbb{E}X_1)(\mathbb{E}_2)$, say X_1 and X_2 are *uncorrelated*

Independence $\implies$ pairwise independent $\implies$ uncorrelated

Theorem 3.7 (L^2 weak law of large numbers). *Given $X_i, i \geq 1$, suppose* $\sup \mathbb{E}X_i^2 \leq c$ *for some constant c and suppose uncorrelatd. Write $\mu_i = \mathbb{E}X_i$, $S_n = \sum_{i=1}^n X_i$, $\bar{\mu}_i = \frac{1}{n}\sum_{i=1}^n \mu_i$.*

Then $\frac{S_n}{n} - \bar{\mu}_n \to 0$ in L^2 as $n \to \infty$.

Proof.

$$\frac{1}{n}\mathbb{E}S_n = \bar{\mu}_n \tag{3.5}$$

$$\mathrm{Var}(S_n) = \sum_{i=1}^{n} \mathrm{Var}(X_i) \leq cn \tag{3.6}$$

$$\mathrm{Var}(\frac{1}{n}S_n) \leq \frac{c}{n} \tag{3.7}$$

$$\mathbb{E}\left(\frac{S_n}{n} - \bar{\mu}_n\right)^2 = \mathrm{Var}\left(\frac{S_n}{n}\right) \leq \frac{c}{n} \to 0 \text{ as } n \to \infty \tag{3.8}$$

□