Independence of random variables

Def Let X be a r.v. The smallest σ-algebra that makes X measurable is

$$\sigma(X) := \{ \underbrace{X^{-1}(B)}_{\substack{\shortparallel \\ \{X\in B\}}} : B \in \mathcal{B} \} = \sigma\left(\{X \le x\} : x \in \mathbb{R} \right)$$

It is called the σ-algebra generated by X

$\{X \le x\} = X^{-1}(-\infty, x]$
and $(-\infty, x]$ generate $\mathcal{B}$

HW?

Def Random variables $X_1, X_2, \ldots$ are independent if $\sigma(X_1), \sigma(X_2), \ldots$ are independent

$$\Updownarrow$$

events $\{X_1 \in B_1\}, \{X_2 \in B_2\}, \ldots$ are independent $\forall\, B_i \in \mathcal{B}$

$$\Updownarrow$$

$$P\{ X_i \in B_i \;\; \forall i \in I \} = \prod_{i \in I} P\{X_i \in B_i\} \qquad \forall \text{ finite } I \subset \mathbb{N}$$

THM R.v's $X_1, X_2, \ldots$ are independent $\iff$

$$P\{ X_i \le x_i \;\; \forall i \in I \} = \prod_{i \in I} P\{X_i \le x_i\} \qquad \forall\, I \subset \mathbb{N},\; \forall\, x_i \in \mathbb{R}$$

Proof $(\Rightarrow)$ by def

$(\Leftarrow)$ Families $\mathcal{A}_i := \{ \{X_i \le x\} : x \in \mathbb{R} \}$ are independent by assumption and are π-systems

$$\Big\Downarrow \text{ Thm p.36}$$

σ-algebras $\sigma(\mathcal{A}_i) = \{ \{X_i \in B\} : B \in \mathcal{R} \}$ are independent. $\square$

Prop $X_1, \ldots X_n$ are independent $\Rightarrow$ $g_1(X_1), \ldots, g_n(X_n)$ are independent $\forall$ measurable functions $g_i : \mathbb{R} \to \mathbb{R}$

$\ulcorner\; \forall B_i \in \mathcal{B}$: $\quad P\{ g_i(X_i) \in B_i \;\; \forall i \in I \} = P\{ X_i \in \underbrace{g_i^{-1}(B_i)}_{\substack{\cap \\ \mathcal{B} \text{ by measurability}}} \;\; \forall i \in I \} \overset{\text{independence}}{=} \prod_{i \in I} P\{X_i \in g_i^{-1}(B_i)\}$

$$= \prod_{i \in I} P\{ g(X_i) \in B_i \}. \quad \square \qquad \lrcorner$$

JOINT DISTRIBUTIONS

- Consider random variables $X_1, \dots, X_n$, possibly dependent.
  Convenient to consider as random vector
  $$X = (X_1, \dots, X_n) \text{ taking values in } \mathbb{R}^n$$
- Similarly to the 1D case, the distribution ("law") of X is
  $$\mathcal{L}_X(B) := \mathbb{P}\{X \in B\}, \quad B \in \mathcal{B}^n$$
  a prob. measure on $(\mathbb{R}^n, \mathcal{B}^n)$.
  $\mathcal{L}_X$ is called the joint distribution of $X_1, \dots, X_n$, as opposed to
  $\mathcal{L}_{X_i}$ = "marginal distribution" of each X_i

Prop $X_1, \dots, X_n$ are independent $\iff \mathcal{L}_X = \mathcal{L}_{X_1} \times \dots \times \mathcal{L}_{X_n}$

For $n=2$.

$(\Rightarrow)$ Let $B = B_1 \times B_2$ where $B_i \in \mathcal{B}$. Then
$$\mathcal{L}_X(B) = \mathbb{P}\{(X_1, X_2) \in B_1 \times B_2\} = \mathbb{P}\{X_1 \in B_1, X_2 \in B_2\}$$
$$= \mathbb{P}\{X_1 \in B_1\} \cdot \mathbb{P}\{X_2 \in B_2\} = \mathcal{L}_{X_1}(B_1) \cdot \mathcal{L}_{X_2}(B_2).$$
$\Rightarrow \mathcal{L}_X$ agrees with $\mathcal{L}_{X_1} \times \mathcal{L}_{X_2}$ on $\underbrace{\text{product sets}}_{\substack{\| \\ \pi\text{-system}}}$ in $\mathcal{B}^2$

$\xrightarrow[\text{p.36}]{\text{Uniqueness}}$ $\mathcal{L}_X$ agrees with $\mathcal{L}_{X_1} \times \mathcal{L}_{X_2}$ on $\sigma(\text{product sets}) = \mathcal{B}^2$

$(\Leftarrow)$ $\mathbb{P}\{X_1 \le x_1, X_2 \le x_2\} = \mathbb{P}\{(X_1, X_2) \in (-\infty, x_1] \times (-\infty, x_2]\}$
$$= \mathcal{L}_X\big((-\infty, x_1] \times (-\infty, x_2]\big) \overset{\text{assm}}{=} \mathcal{L}_{X_1}(-\infty, x_1] \cdot \mathcal{L}_{X_2}(-\infty, x_2] = \mathbb{P}\{X_1 \le x_1\} \cdot \mathbb{P}\{X_2 \le x_2\}$$
$\xrightarrow[\text{p.37}]{\text{criterion}}$ $X_1 \perp\!\!\!\perp X_2$.

Def Joint CDF of $X_1, \dots, X_n$ is $F_X(x_1, \dots, x_n) := \mathbb{P}\{X_1 \le x_1\} \cdots \mathbb{P}\{X_n \le x_n\}$

Criterion p.37 $\Rightarrow$

Cor $X_1, \dots, X_n$ are independent $\iff F_X(x_1, \dots, x_n) = F_{X_1}(x_1) \cdots F_{X_n}(x_n) \quad \forall x_i \in \mathbb{R}$

Expectation : Similarly to 1-D case :

$$\mathbb{E}g(X) = \int_{\mathbb{R}^n} g \, d\mathcal{L}_X \,, \quad \text{i.e.} \quad \mathbb{E}g(X_1,\dots,X_n) = \int_{\mathbb{R}^n} g(X_1,\dots,X_n)\, d\mathcal{L}_X(x_1,\dots,x_n)$$

THM $X_1,\dots,X_n$ are independent $\Rightarrow$ $\mathbb{E}[X_1 \cdots X_n] = (\mathbb{E}X_1)\cdots(\mathbb{E}X_n)$

For $n=2$:

$$\mathbb{E}[X_1X_2] = \int_{\mathbb{R}^2} x_1x_2 \, d\overset{\mathcal{L}_{X_1}\times\mathcal{L}_{X_2}\ (\text{p. }38)}{\overset{\parallel}{\mathcal{L}_X}} \overset{}{\underset{\text{Fubini}}{=\!=\!=}} \left(\int_{\mathbb{R}} x_1 \, d\mathcal{L}_{X_1}\right)\left(\int_{\mathbb{R}} x_2 \, d\mathcal{L}_{X_2}\right) = (\mathbb{E}X_1)(\mathbb{E}X_2)$$

Remarks (a) "$\Leftarrow$" does not hold, i.e. if $X_1 = X_2$: $\mathbb{E}X^2 = (\mathbb{E}X)^2$ would mean that $Var(X)=0$.

(b) If X_i are independent, then so are $g_i(X_i)$ $\forall$ mble g_i (p. 37)

$$\Rightarrow \mathbb{E}[g_1(X_1)\cdots g_n(X_n)] = (\mathbb{E}g_1(X_1))\cdots(\mathbb{E}g_n(X_n)) \quad \forall \text{ mble } g_i \qquad (*)$$

(c) $(*) \Rightarrow X_1,\dots,X_n$ are independent

Choose $g_i(x) = \mathbb{1}_{\{x \le x_i\}}$

$\Rightarrow$ $(*)$ becomes $\mathbb{P}\{X_1 \le x_1,\dots,X_n \le x_n\} = \mathbb{P}\{X_1\le x_1\}\cdots\mathbb{P}\{X_n \le x_n\}$

$\Rightarrow X_1,\dots X_n$ are independent (p. 38)

Cor $X_1,\dots,X_n$ are independent $\Rightarrow$

$$Var\left(\sum_{i=1}^n X_i\right) = \sum_{i=1}^n Var(X_i)$$

Proof W.L.O.G. $\mathbb{E}X_i = 0$ (subtracting $\mathbb{E}X_i$ does not affect the variance)

$$Var\left(\sum_{i=1}^n X_i\right) = \mathbb{E}\left[\left(\sum_{i=1}^n X_i\right)^2\right] = \mathbb{E}\left[\sum_{i,j=1}^n X_iX_j\right] = \sum_{i,j=1}^n \underbrace{\mathbb{E}[X_iX_j]}_{\substack{\mathbb{E}[X_i^2] = Var(X_i) \text{ if } i=j \\ \mathbb{E}[X_i]\mathbb{E}[X_j] = 0\cdot 0 = 0 \text{ if } i\ne j}} = \sum_{i=1}^n Var(X_i).$$

Ex: $X \sim Binom(n,p)$ = # successes in n independent trials, where each trial is a success with probability $= p$.

$X = \sum_{i=1}^n X_i$, $X_i \sim Ber(p)$ iid. $\Rightarrow$ $\mathbb{E}X = np$, $Var(X) = np(1-p)$.