Beginning of Sept. 14, 2022

Theorem

Let $\{X_{i,j} : i \leqslant n, j \leqslant m(i)\}$ be independent. Then $f(X_{i,1}, \ldots, X_{i,m(i)})$, $i \leqslant n$, are independent random variables.

Proof. Let $\mathcal{F}_{i,j} = \sigma(X_{i,j})$ and $\mathcal{B}_i := \sigma(\mathcal{F}_{i,1}, \ldots, \mathcal{F}_{i,m(i)})$. By the previous theorem each $\mathcal{B}_i$'s are independent. Each f_i is $\mathcal{B}_i$-measurable so the random variables f_i are independent. □

Fubini theorem says for $f(x, y)$ on $\Omega \times \Omega_2$,

$$\int f \, \mathrm{d}(\mu_1 \times \mu_2) = \int_{\Omega_1} \int_{\Omega_2} f \, \mathrm{d}\mu_2 \, \mathrm{d}\mu_1$$

provided $f \geqslant 0$ or f is integrable (i.e., $\int |f| \, \mathrm{d}(\mu_1 \times \mu_2) < \infty$). (Here since μ_i's are probability measures they are assumed to be σ-finite.) For random variables:

Theorem: D2.1.12

Let X, Y be independent with distributions μ_X and μ_Y on $\mathbb{R}$. Let $h : \mathbb{R}^2 \to \mathbb{R}$ satisfy either $h \geqslant 0$ or $\mathbb{E}|h(X, Y)| < \infty$. Then

$$\mathbb{E}h(X, Y) = \iint_{\mathbb{R}^2} h(X, Y) \, \mathrm{d}\mu_X(\mathrm{d}x)\mu_Y(\mathrm{d}y) \tag{*}$$

and the other of integration does not matter.
In particular, for $h(x, y) = f(x)g(y)$ with either $f, g \geqslant 0$ or $\mathbb{E}|f(X)|, \mathbb{E}|g(Y)| < \infty$, we have

$$\mathbb{E}[f(X)g(Y)] = \mathbb{E}f(X)\mathbb{E}g(Y), \tag{**}$$

i.e., independence $\Rightarrow$ (product of $\mathbb{E}$ = $\mathbb{E}$ of product).

Proof. (*) follows from Fubini since the distribution of (X, Y) is $\mu_X \times \mu_Y$ by independence.
For (**),

$$\begin{aligned}
\mathbb{E}[f(X)g(Y)] &= \iint_{\mathbb{R}^2} f(x)g(y) \, \mu_X(\mathrm{d}x)\mu_Y(\mathrm{d}y) \\
&= \int_{\mathbb{R}} g(y) \int_{\mathbb{R}} f(x)\mu_X(\mathrm{d}x)\mu_Y(\mathrm{d}y) \\
&= \left(\int_{\mathbb{R}} f(x)\mu_X(\mathrm{d}x) \right) \int_{\mathbb{R}} g(y)\mu_Y(\mathrm{d}y) = \mathbb{E}f(X)\mathbb{E}g(Y).
\end{aligned}$$

□

By induction, we may generalize the above result into any finite number of random variables. That is, for independent $X_1, \ldots, X_n$ with $\mathbb{E}|\prod X_i| < \infty$, we have $\mathbb{E}[\prod X_i] = \prod \mathbb{E}X_i$.

Sums of Independent Random Variables

Let X, Y be independent with distribution functions F and G. The d.f. of $X + Y$ is the **convolution**

$$H(z) = \mathbb{P}(X + Y \leqslant z) = \int_{-\infty}^{\infty} F(z - y)\mathrm{d}(G(y)) =: (F * G)(z).$$

To see this, we apply Fubini to $1_{x+y\leqslant z}$:

$$H(z) = \mathbb{E}1_{\{X+Y\leqslant z\}} = \iint 1_{\{x\leqslant z-y\}}\, \mathrm{d}F(x)\mathrm{d}G(y) = \int F(z-y)\, \mathrm{d}G(y).$$

Note by doing $\{y \leqslant x - z\}$ first we obtain see that $F * G \equiv G * F$.

Example. Let X be uniform on $[0,2]$ and Y is exponential with parameter λ. That is, X has $1/2$ on $[0,2]$ and Y has $\lambda e^{-\lambda y}$ on $[0,\infty)$. The distribution function of Y is $1-e^{-\lambda y}$ for $y \geqslant 0$. Then

$$H(z) = \int_{-\infty}^{\infty} \underbrace{(1-e^{-\lambda(z-y)})1_{\{z-y\geqslant 0\}}}_{F(z-y)} \underbrace{\frac{1}{2}1_{[0,2]}\,\mathrm{d}y}_{\mathrm{d}G(y)}.$$

That is,

$$H(z) = \begin{cases} 0 & z<0 \\ \displaystyle\int_0^z (1-e^{-\lambda z}e^{\lambda y})/2\,\mathrm{d}y & 0\leqslant z\leqslant 2 \\ \displaystyle\int_0^2 (1-e^{-\lambda z}e^{\lambda y})/2\,\mathrm{d}y & z>2 \end{cases} = \begin{cases} 0 & z<0 \\ \dfrac{z}{2} - \dfrac{1-e^{-\lambda z}}{2\lambda} & 0\leqslant z\leqslant 2 \\ 1-e^{-\lambda z}\dfrac{e^{2\lambda}-1}{2\lambda} & z>2. \end{cases}$$

Let μ_n be a probability measure on $(\mathbb{R}^n, \mathcal{R}^n)$. We can make a random vector with distribution μ_n:we take $(\Omega, \mathcal{F}, \mathbb{P}) = (\mathbb{R}^n, \mathcal{R}^n, \mu_n)$ and $(X_1, \ldots, X_n)$ to be identity.

Infinite Sequence of Random Variables

We say **finite-dimensional** distributions of $\{X_n, n \geqslant 1\}$ are all distributions of form $\{X_i, i \leqslant I\}$ for $I \subset \mathbb{N}$ finite. By using marginals is sufficient to consider $I = \{1, \ldots, n\}$. Suppose we are given μ_n on $(\mathbb{R}^n, \mathcal{R}^n)$ for every n.
Question: is there a $\mathbb{P}$ on $(\mathbb{R}^{\mathbb{N}}, \mathcal{R}^{\mathbb{N}})$ with distribution μ_n for the first n coordinates? That is,

$$\mathbb{P}(A_1 \times \ldots \times A_n \times \mathbb{R} \times \ldots) = \mu_n(A_1 \times \ldots \times A_n)?$$

(Well of course no, since if $n > m$, μ_n determines what μ_m would be.) What if this consistency is satisfied?

Theorem: Kolmogorov Extension Theorem

Let μ_n be a p.m. on $(\mathbb{R}^n, \mathcal{R}^n)$ for all n. Suppose consistency holds among the μ_n's, i.e.,

$$\mu_{n+1}(A \times \mathbb{R}) = \mu_n(A) \qquad \text{for all } A = \prod_{i=1}^n (a_i, b_i] \text{ and } n \geqslant 1.$$

(In reality the above choice of A can be anything in $\mathcal{R}^n$; we just picked the most canonical one.) Then there exists a unique $\mathbb{P}$ on $(\mathbb{R}^{\mathbb{N}}, \mathcal{R}^{\mathbb{N}})$ with

$$\mathbb{P}(\prod_{i=1}^n (a_i, b_i] \times \mathbb{R} \times \mathbb{R} \times \ldots) = \mu_n(\prod_{i=1}^n (a_i, b_i]).$$

(Sets of form $A \times \mathbb{R} \times \mathbb{R} \times \ldots$ with $A \in \mathcal{R}^n$ is a **cylinder set** in $\mathbb{R}^{\mathbb{N}}$. They form a σ-field.)

0.1 Weak Laws of Large Numbers

Some recaps:

- We say $X_n \to X$ **in probability** (i.e. in measure) if $\mathbb{P}(|X_n - X| > \epsilon) \to 0$ as $n \to \infty$.
- We say $X_n \to X$ **in L^p** (where $p > 0$) if $\mathbb{E}(|X_n - X|^p) \to 0$ as $n \to \infty$.
  - For $p \geqslant 1$, this is equivalent to $\|X_n - X\|_p \to 0$.
  - For $p < 1$ this does not hold as such $\|\cdot\|_p$ does not define a norm.
  - In principle the X_n can have infinite p^{th} moment but the definition still makes sense.