L^2 [§10.5 in Driver]

Given a measure space $(\Omega, \mathcal{F}, \mu)$,

$L^2(\Omega, \mathcal{F}, \mu) = \{ f : \Omega \to \mathbb{R}$ measurable s.t. $\int_\Omega f^2 \, d\mu < \infty \}$.

Note: for any real numbers f, g,

$$0 \le (|f| - |g|)^2 = f^2 - 2|fg| + g^2$$

$$\therefore \quad |fg| \le \frac{1}{2}(f^2 + g^2)$$

Thus, if $f, g \in L^2$, then $\int |fg| \, d\mu \le \int \frac{1}{2}(f^2 + g^2) \, d\mu$

$$= \frac{1}{2} \int f^2 \, d\mu + \frac{1}{2} \int g^2 \, d\mu < \infty.$$

I.e. **Prop**: If $f, g \in L^2$ then $fg \in L^1$.

Cor: If μ is a finite measure, then $L^2(\mu) \subseteq L^1(\mu)$

B/c $g = 1 \in L^2$: $\int 1^2 \, d\mu = \mu(\Omega) < \infty$ $\therefore$ If $f \in L^2$, $f \cdot 1 \in L^1$.

Cauchy-Schwarz

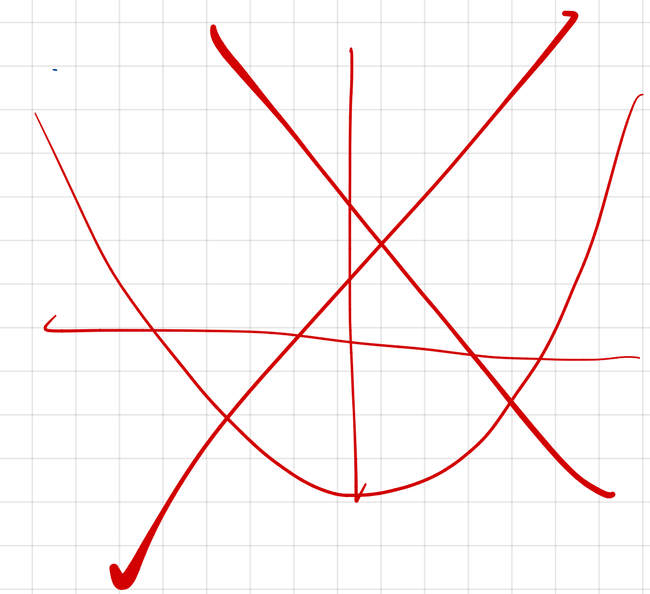
For $f, g \in L^2(\Omega, \mathcal{F}, \mu)$, define:

$$\|f\|_{L^2} := \left(\int_\Omega f^2 \, d\mu\right)^{1/2}, \qquad \langle f, g \rangle_{L^2} := \int_\Omega fg \, d\mu$$

Theorem: If $f, g \in L^2(\Omega, \mathcal{F}, \mu)$, then

$$|\langle f, g \rangle_{L^2}| \le \int |fg| \, d\mu \le \|f\|_{L^2} \|g\|_{L^2}$$

$|\int fg| \le \int |fg|$ ✓



Pf.

For $t \in \mathbb{R}$, $p(t) = \int (|f| - t|g|)^2 \, d\mu \ge 0$

$$= \int (f^2 - 2t|fg| + t^2 g^2) \, d\mu$$

$$= \int f^2 \, d\mu - 2\int |fg| \, d\mu \cdot t + \int g^2 \, d\mu \cdot t^2$$

$$= \|f\|_{L^2}^2 - 2\int |fg| \, d\mu \cdot t + \|g\|_{L^2}^2 \, t^2$$

$$\left(-2\int |fg| \, d\mu\right)^2 - 4\|f\|_{L^2}^2 \|g\|_{L^2}^2 \le 0.$$

Cor: $L^2(\Omega, \mathcal{F}, \mu)$ is a vector space, and $\|f\|_{L^2}^2 = \sqrt{\langle f, f\rangle_{L^2}}^2$ is a norm on it.

$\int f^2 d\mu = \int f \cdot f d\mu$ ✓

Pf. If $\alpha \in \mathbb{R}$, $f \in L^2$,

$$\|\alpha f\|_{L^2}^2 = \int (\alpha f)^2 d\mu = \alpha^2 \int f^2 d\mu = \alpha^2 \|f\|_{L^2}^2$$

$$\therefore \|\alpha f\|_{L^2} = |\alpha| \|f\|_{L^2}$$

If $f, g \in L^2$, $\|f+g\|_{L^2}^2 = \int (f+g)^2 d\mu = \int (f^2 + 2fg + g^2) d\mu$

$$= \int f^2 + 2\int fg + \int g^2$$

$$\leq \int f^2 + 2\|f\|_{L^2}\|g\|_{L^2} + \int g^2$$

$$= \|f\|_{L^2}^2 + 2\|f\|_{L^2}\|g\|_{L^2} + \|g\|_{L^2}^2$$

$$= (\|f\|_{L^2} + \|g\|_{L^2})^2 < \infty$$

If $\|f\|_{L^2} = 0$

then $\int f^2 d\mu = 0$

$\Rightarrow f^2 = 0$ a.s. $[\mu]$

$\Rightarrow f = 0$ a.s. $[\mu]$.

///

Fact: L^2 is actually a Hilbert space: it is (Cauchy) complete.

I.e. If $f_n \in L^2$ s.t. $\|f_n - f_m\|_{L^2} \to 0$ as $n, m \to \infty$,

then $\exists!\, f \in L^2$ s.t. $\|f_n - f\|_{L^2} \to 0$.

(The same holds true in L^1, and in general in L^p for $1 \le p \le \infty$.)

We'll return to prove this important fact in the near future.

This will be very important when we make a serious study of conditional probability.

Covariance

In a probability space $(\Omega, \mathcal{F}, \mathbb{P})$, $L^2 \subseteq L^1$. In fact

$$\|X\|_{L^1} = \int_\Omega |X| \, d\mathbb{P} = \int_\Omega |X \cdot 1| \, d\mathbb{P} \underset{C-S}{\leq} \|X\|_{L^2} \|1\|_{L^2} = \|X\|_{L^2}$$

$$\int_\Omega 1^2 \, d\mathbb{P} = 1$$

I.e. $|\mathbb{E}[X]| \leq \mathbb{E}[|X|] \leq \sqrt{\mathbb{E}[X^2]}$

Def For $X, Y \in L^2$, let $\mathring{X} = X - \mathbb{E}[X]$, $\mathring{Y} = Y - \mathbb{E}[Y]$ "centered"

$$\mathbb{E}[\mathring{X}] = \mathbb{E}[X - \mathbb{E}[X]] = \mathbb{E}[X] - \mathbb{E}[X] = 0$$

Their covariance is

$$\begin{aligned}
\mathrm{Cov}(X,Y) &:= \mathbb{E}[\mathring{X}\mathring{Y}] \\
&= \mathbb{E}[(X - \mathbb{E}[X])(Y - \mathbb{E}[X])] \\
&= \mathbb{E}[XY - X\mathbb{E}[Y] - \mathbb{E}[X]Y + \mathbb{E}[X]\mathbb{E}[Y]] \\
&= \mathbb{E}[XY] - \mathbb{E}[X]\mathbb{E}[Y].
\end{aligned}$$

For $X \in L^2$, its variance is

$$\mathrm{Var}(X) := \mathrm{Cov}(X,X) = \mathbb{E}[\mathring{X}^2] = \mathbb{E}[X^2] - (\mathbb{E}[X])^2 \geq 0$$

Lemma: If $X \in L^2(\Omega, \mathcal{F}, P)$ and $Var(X) = 0$, then $X \equiv$ Const. a.s.

Pf. $0 = Var(X) = E[\mathring{X}^2] \Rightarrow \mathring{X}^2 = 0$ a.s. $\Rightarrow \mathring{X} = 0$ a.s.

$\mathring{X}^2 \geq 0$ $\qquad$ $\overset{\shortparallel}{X - E[X]}$ $\therefore X = E[X]$ a.s.

///

Eg. $X \stackrel{d}{=} N(\alpha, t)$. $\quad X \stackrel{d}{=} \sqrt{t} Z + \alpha$, $Z \stackrel{d}{=} N(0,1)$

$$Var(X) = E[(X-\alpha)^2] = E[(\sqrt{t}Z)^2] = t\, E[\underbrace{Z^2}_{1}] = t.$$

Eg. $T \stackrel{d}{=} Exp(\alpha)$, $\quad E[T] = \frac{1}{\alpha}$

$$Var(T) = E[(T - \tfrac{1}{\alpha})^2] = \int_0^\infty (t - \tfrac{1}{\alpha})^2 \alpha e^{-\alpha t} dt = \frac{1}{\alpha^2}$$

Eg. $N \stackrel{d}{=} Pois(\alpha)$, $\quad E[N] = \alpha$

$$Var(N) = E[(N-\alpha)^2] = \sum_{k=0}^{\infty} (k-\alpha)^2 e^{-\alpha} \frac{\alpha^k}{k!} = \alpha$$

Def: $X, Y \in L^2(\Omega, \mathcal{F}, \mathbb{P})$ are uncorrelated if $\mathrm{Cov}(X,Y) = 0$.

In general, their correlation is $\mathrm{Corr}(X,Y) = \dfrac{\mathrm{Cov}(X,Y)}{\sqrt{\mathrm{Var}(X)\,\mathrm{Var}(Y)}}$

$= \left| \dfrac{\mathbb{E}[\mathring{X}\mathring{Y}]}{\|\mathring{X}\|_{L^2} \|\mathring{Y}\|_{L^2}} \right| \le 1$ (C-S)

$\|\mathring{X}\|_{L^2}$ = Standard Deviation $= \sqrt{\mathrm{Var}(X)}$

Prop: $\mathrm{Cov}(X+\alpha, Y) = \mathrm{Cov}(X, Y+\alpha) = \mathrm{Cov}(X,Y)$ for any $\alpha \in \mathbb{R}$. As a result, if $X_1, \dots, X_n$ are all (pairwise) uncorrelated, then

$$\mathrm{Var}(X_1 + \cdots + X_n) = \mathrm{Var}(X_1) + \cdots + \mathrm{Var}(X_n).$$

Pf. $(X+\alpha)^\circ = \mathring{X}$

$$\mathrm{Var}(X_1 + \cdots + X_n) = \mathbb{E}\left[(\mathring{X}_1 + \cdots + \mathring{X}_n)^2\right]$$

$$= \mathbb{E}\left[\sum_j \mathring{X}_j \cdot \sum_k \mathring{X}_k\right] = \sum_{j,k=1}^n \mathbb{E}[\mathring{X}_j \mathring{X}_k] = \sum_{j,k=1}^n \underbrace{\mathrm{Cov}(X_j, X_k)}_{=0 \text{ if } j \ne k}$$

$$= \sum_{j=1}^n \mathrm{Cov}(X_j, X_j)$$

$$= \sum_j \mathrm{Var}(X_j) \quad ///$$

What Does "Uncorrelated" Mean?

Eg. Let $A, B \in \mathcal{F}$, and consider $\mathbb{1}_A, \mathbb{1}_B$

$X = \mathbb{1}_A$

$$\left.\begin{aligned} P(X=1) &= P(A) \\ P(X=0) &= P(A^c) = 1 - P(A) \end{aligned}\right\} \text{Bernoulli}$$

$$\begin{aligned} &\mathrm{Cov}(\mathbb{1}_A, \mathbb{1}_B) \\ &= \mathbb{E}(\mathbb{1}_A \mathbb{1}_B) - \mathbb{E}(\mathbb{1}_A)\mathbb{E}(\mathbb{1}_B) \\ &= \mathbb{E}(\mathbb{1}_{A\cap B}) - \mathbb{E}(\mathbb{1}_A)\mathbb{E}(\mathbb{1}_B) = P(A\cap B) - P(A)P(B) = 0 \text{ iff } A, B \text{ are independent} \end{aligned}$$

Eg. Toss a fair coin n times. $X_j = \begin{cases} 1 & \text{if the } j^{th} \text{ toss is Heads} \\ 0 & \text{" Tails} \end{cases}$

$$\Omega = \{ (\omega_1, \dots, \omega_n) : \omega_j \in \{0, 1\} \} \quad \hookrightarrow X_j(\omega) = \omega_j$$

$$\mathcal{F} = 2^{\Omega}$$

$$\mathbb{P}(A) = \frac{\# A}{2^n}$$

Exercise: $\mathrm{Cov}(X_j, X_k) = \delta_{jk}$