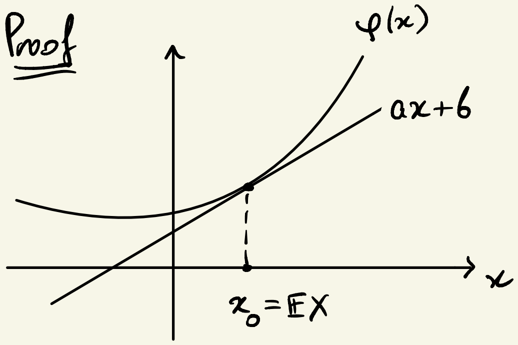
INEQUALITIES

THM (Jensen's inequality) If $\varphi : \mathbb{R} \to \mathbb{R}$ is a convex function then
$$\varphi(\mathbb{E}X) \leq \mathbb{E}\varphi(X)$$

Proof



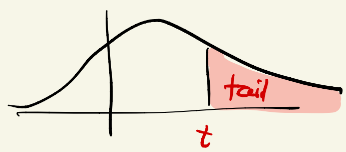
Let $ax+b$ be a tangent to $\varphi(x)$ at $x_0 = \mathbb{E}X$
$\Rightarrow \quad \varphi(x) \geq ax+b \quad \forall x \in \mathbb{R}$.
Substitute $x = X$, take expectation $\Rightarrow$
$\mathbb{E}\varphi(X) \geq a\,\mathbb{E}X + b = ax_0 + b = \varphi(x_0) = \varphi(\mathbb{E}X)$. $\square$

strict ineq. $\to$ KW ?

Ex: $\|X\|_{L^p} := \left(\mathbb{E}|X|^p\right)^{1/p}$ increase in p

[To check $\|X\|_{L^p} \leq \|X\|_{L^q}$ for $p<q$, apply Jensen for $\varphi(x) = |x|^{q/p}$]

THM (Markov's inequality) $\forall$ r.v. $X \geq 0$:
$$\mathbb{P}\{X \geq t\} \leq \frac{\mathbb{E}X}{t}$$



Proof $\quad x \geq t\,\mathbb{1}_{\{x \geq t\}} \quad \forall x, t \geq 0$.

Apply this for X and take expectations on both sides:
$$\mathbb{E}X \geq t\,\mathbb{P}\{X \geq t\}. \quad \square$$

- **Prop** (Chebyshev's inequality) $\forall$ r.v. X with mean μ, variance σ^2:

$$P\{|X-\mu| \geq t\} \leq \frac{\sigma^2}{t^2} \qquad \forall t>0$$

Proof $P\{|X-\mu|\geq t\} = P\{(X-\mu)^2 \geq t^2\} \leq \frac{\mathbb{E}(X-\mu)^2}{t^2}$ (Markov for $(X-\mu)^2$)

$= \sigma^2/t^2$. QED

Ex: With prob ≥ 0.88, at most 3 students get their own exams $\left(\begin{array}{l}\mu=\sigma^2=1 \text{ p.31} \\ t=1/3\end{array}\right)$

OPTIMALITY

- Markov's inequality is best possible if one only knows $\mathbb{E}X$
- Chebyshev's is best possible if one only knows $\mathbb{E}X$, $\mathrm{Var}(X)$.

HW

Ex $X \sim N(0,1)$: Cheb $\Rightarrow$ $P\{|X|>t\} \leq \frac{1}{t^2}$

much weaker than the Gaussian tail bound (p.28!) ☹

However, a trick: fix $\forall \lambda>0$;

$$P\{X>t\} = P\{e^{\lambda X} > e^{\lambda t}\} \overset{\text{Markov}}{\leq} e^{-\lambda t}\, \underbrace{\mathbb{E}\left[e^{\lambda X}\right]}_{= \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{\infty} e^{\lambda x} e^{-x^2/2}\,dx \,=\, e^{\lambda^2/2}}$$

(check! MGF)

$$= \exp(-\lambda t + \lambda^2/2)$$

optimize $\lambda = t$

$$= \exp(-t^2/2) \quad \leftarrow \text{Gaussian tail bound!} ☺$$

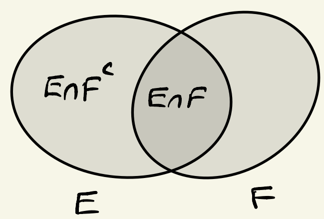
INDEPENDENCE

Def Events E, F are independent $(E \perp\!\!\!\perp F)$ if
$$\mathbb{P}(E \cap F) = \mathbb{P}(E) \cdot \mathbb{P}(F)$$

Ex Flip a coin twice, $E := \{H \text{ of } 1^{st} \text{ flip}\}$, $F = \{H \text{ on } 2^{nd} \text{ flip}\}$
$$\mathbb{P}(E \cap F) = \mathbb{P}(\{HH\}) = \tfrac{1}{4} \,; \quad \mathbb{P}(E) = \mathbb{P}(F) = \tfrac{1}{2} \quad \Rightarrow \quad E \perp\!\!\!\perp F.$$

Prop (Stability) $E \perp\!\!\!\perp F \Rightarrow E^c \perp\!\!\!\perp F,\ E \perp\!\!\!\perp F^c,\ E^c \perp\!\!\!\perp F^c$

Proof: $E \perp\!\!\!\perp F^c$?



$$\mathbb{P}(E) = \underbrace{\mathbb{P}(E \cap F)}_{\substack{\| \\ \mathbb{P}(E)\,\mathbb{P}(F) \text{ by assumption}}} + \mathbb{P}(E \cap F^c)$$
$$\Rightarrow \mathbb{P}(E \cap F^c) = \mathbb{P}(E)\big(1 - \mathbb{P}(F)\big) = \mathbb{P}(E) \cdot \mathbb{P}(F^c). \qquad \square$$

Note: $E \perp\!\!\!\perp \emptyset$, and thus $E \perp\!\!\!\perp \Omega$, for any event E.

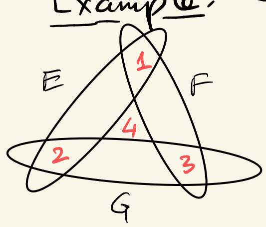
Warning: disjoint events are not independent (unless $= \emptyset$ or Ω)

Independence of more than 2 events

Def Events $E_1, E_2, \dots$ are independent if
$$\mathbb{P}\Big(\bigcap_{i \in I} E_i\Big) = \prod_{i \in I} \mathbb{P}(E_i) \qquad \forall \text{ finite } I \subset \mathbb{N}$$

Warning: pairwise independence $(E_i \cap E_j = \emptyset\ \forall i, j) \not\Rightarrow$ independence

Example: $\Omega = \{4 \text{ points}\}$, $\mathbb{P} = \text{unif}(\Omega)$



$$\mathbb{P}(E) = \mathbb{P}(F) = \mathbb{P}(G) = \tfrac{2}{4} = \tfrac{1}{2} \,;$$
$$\mathbb{P}(\underbrace{E \cap F}_{\substack{\| \\ \{1\}}}) = \frac{1}{4} = \mathbb{P}(E)\,\mathbb{P}(F) \Rightarrow E \perp\!\!\!\perp F.$$

Independence is a strong property.
Requires 2^n equations to hold

Similarly for the other pairs $\Rightarrow$ $\boxed{\text{pairwise indep}}$

But $\mathbb{P}(\underbrace{E \cap F \cap G}_{\substack{\| \\ \emptyset}}) = 0 \neq \mathbb{P}(E)\,\mathbb{P}(F)\,\mathbb{P}(G) \Rightarrow \boxed{\text{NOT indep}}$

Def Families of events $\mathcal{E}_1, \mathcal{E}_2, \dots$ are independent if events $E_1, E_2, \dots$ are independent $\forall E_i \in \mathcal{E}_i$

Warning: $\mathcal{E} \perp\!\!\!\perp \mathcal{F} \not\Rightarrow \sigma(\mathcal{E}) \perp\!\!\!\perp \sigma(\mathcal{F})$ ☹

σ-algebras generated by $\mathcal{E}, \mathcal{F}$

Consider $\mathcal{E} := \{E, G\}$, $\mathcal{F} = \{F\}$ in the example p. 34.

Then $\mathcal{E} \perp\!\!\!\perp \mathcal{F}$ by pairwise independence,

but $E \cap G = \{2\}$ and $G = \{1,3\}$ are disjoint $\Rightarrow \underset{\sigma(\mathcal{E})}{E \cap G} \not\perp\!\!\!\perp \underset{\sigma(\mathcal{F})}{F} \Rightarrow \sigma(\mathcal{E}) \not\perp\!\!\!\perp \sigma(\mathcal{F})$

- WILL SHOW: if each family $\mathcal{E}_i$ is closed under finite intersections, then $\sigma(\mathcal{E}_i)$ are independent. ☺ ($\mathcal{E}$ would be forced to contain $E \cap G$)

This follows from Dynkin's π–λ theorem

Def A family of sets $\mathcal{P}$ is called a π-system if A is closed under finite intersections, i.e.

$$E, F \in \mathcal{P} \Rightarrow E \cap F \in \mathcal{P}.$$

Ex: {rectangles} in $\mathbb{R}^2$

Def A family of sets $\mathcal{L}$ is called a λ-system if $\mathcal{L}$ is closed under complements and countable disjoint unions, i.e.

(i) $E \in \mathcal{L} \Rightarrow E^c \in \mathcal{L}$

(ii) $E_1, E_2, \dots \in \mathcal{L}$ are disjoint $\Rightarrow \bigsqcup_{i=1}^{\infty} E_i \in \mathcal{L}$

- Remark: The notion of λ-system is strictly weaker than σ-algebra:

Example All events in $(\Omega, \Sigma, \mathbb{P})$ that are independent of a given event E, i.e. $\mathcal{L} := \{E \in \Sigma : E \perp\!\!\!\perp F\}$ is a λ-system. But not always a σ-algebra.

Example: $\mathcal{L} = \{E, G\}$ in ex. above.

λ-system?

(i) $E \perp\!\!\!\perp F \Rightarrow E^c \perp\!\!\!\perp F$ (stability p. 34)

(ii) $E_i \perp\!\!\!\perp F$ disjoint $\Rightarrow \bigsqcup_{i=1}^{\infty} E_i \perp\!\!\!\perp F$?

$$\mathbb{P}\Big(\big(\bigsqcup_i E_i\big) \cap F\Big) = \mathbb{P}\Big(\bigsqcup_i (E_i \cap F)\Big) = \sum_i \mathbb{P}(E_i \cap F) = \sum_i \mathbb{P}(E_i)\mathbb{P}(F)$$

$$= \Big(\sum_i \mathbb{P}(E_i)\Big)\mathbb{P}(F) = \mathbb{P}\big(\bigsqcup_i E_i\big)\mathbb{P}(F) \quad \text{Yes.}$$

THM (Dynkin's π-λ theorem)
Let $\mathcal{P}$ be a π-system and $\mathcal{L}$ be a λ-system. Then
$$\mathcal{P} \subset \mathcal{L} \Rightarrow \sigma(\mathcal{P}) \subset \mathcal{L}.$$

Trivial for σ-algebras $\mathcal{L}$.

THM If $\mathcal{E}_1, \mathcal{E}_2, \dots$ are independent π-systems, then $\sigma(\mathcal{E}_1), \sigma(\mathcal{E}_2), \dots$ are independent.

Proof for two π-systems: (general case $\to$ **HW**)
$$\mathcal{E} \perp\!\!\!\perp \mathcal{F} \overset{?}{\Longrightarrow} \sigma(\mathcal{E}) \perp\!\!\!\perp \sigma(\mathcal{F})$$
Enough to check that $\forall\, F \in \mathcal{F}$:
$$\mathcal{E} \perp\!\!\!\perp F \Rightarrow \sigma(\mathcal{E}) \perp\!\!\!\perp F$$
Assume $\mathcal{E} \perp\!\!\!\perp F$. Consider the family of all events independent of F:
$$\mathcal{L} := \{ E \in \Sigma : E \perp\!\!\!\perp F \}.$$
Then $\mathcal{L}$ is a λ-system (Ex. p. 35), and $\mathcal{E} \subset \mathcal{L}$ by assumption.
$$\xrightarrow{\pi\text{-}\lambda \text{ thm}} \sigma(\mathcal{E}) \subset \mathcal{L} \Rightarrow \sigma(\mathcal{E}) \perp\!\!\!\perp F.$$
$\square$

Another (classical) use of π-λ thm:

THM (Uniqueness of measures)
If measures μ, μ' on $(\Omega, \mathcal{F})$ agree on all sets that form a π-system $\mathcal{P} \subset \mathcal{F}$ then μ, μ' agree on $\sigma(\mathcal{P})$.

e.g. intervals

Borel sets

Proof $\mathcal{L} := \{ B \in \sigma(\mathcal{P}) : \mu(B) = \mu'(B) \}$
is a λ-system (check! easy)
$\mathcal{P} \subset \mathcal{L}$ by assumption. $\xrightarrow{\pi\text{-}\lambda \text{ thm}} \sigma(\mathcal{P}) \subset \mathcal{L} \Rightarrow \mu, \mu'$ agree on $\forall\, B \in \sigma(\mathcal{P})$. $\square$