

Introduction à la théorie de la mesure et de l'intégration

L'objet de ce texte n'est pas de donner des définitions, théorèmes ou démonstrations sur la théorie de la mesure et de l'intégration. Plutôt, il s'agit de fournir du contexte, c'est-à-dire des éléments de réponse aux questions suivantes : quel est le but de cette théorie ? en quoi est-elle utile et digne d'intérêt ? Il me semble bienvenu de répondre à ces questions pour deux raisons. D'une part, y répondre présente un intérêt intrinsèque. D'autre part, disposer d'éléments de réponse à ces questions vous donnera de la motivation pour bien travailler dans cette UE : quand on comprend qu'une théorie a de la valeur, en acquérir la maîtrise devient une perspective attrayante.

1 Une théorie générale des pondérations

Imaginez que vous disposez d'un ensemble E et que vous souhaitez « répartir de la masse » sur E . Par exemple, disons que E désigne une carte du monde et qu'on s'intéresse à la masse de cacao produite au printemps 2025, en assignant chaque cabosse de cacao à son lieu de production et en exprimant les poids en kilogrammes. Comment encoder cela dans un objet mathématique propre ?

On peut pour cela imaginer plusieurs façons de procéder. Par exemple, pour chaque pays, on peut comptabiliser la production totale. Mais comme on souhaite vraiment se représenter la masse comme « répartie sur la carte du monde », il s'agira alors d'encoder les choses par une fonction $f : E \rightarrow [0, \infty[$ qui, pour chaque pays, vaudra partout sur ce pays la masse produite divisée par la superficie du pays exprimée en km^2 — de même que si on étale un volume fixé de confiture de façon régulière sur une tartine, la hauteur de confiture vaut le volume de confiture *divisé par l'aire* de la tartine. Ainsi, si deux pays ont la même production, f sera plus élevée sur le petit pays, rendant compte du fait que la production y est plus « concentrée ».

Mais cette façon de modéliser les choses n'est pas la seule possible. On pourrait aussi considérer que chaque cabosse vient d'un arbre bien précis ayant des coordonnées GPS exactes. Dans ce cas, on serait bien en peine de mimer l'approche précédente, puisqu'un point a une aire nulle. Sur le principe et si on fait abstraction de la lisibilité, il s'agirait plutôt de représenter chaque arbre par un point sur la carte et d'écrire à côté de chaque arbre sa production.

On se heurte alors à un problème : comment développer une théorie permettant de gérer simultanément la première et la seconde approche ? Comment développer une théorie qui permette de gérer à la fois des répartitions de masse « continues » (étalées) et des répartitions discrètes (concentrées sur des points) ?

La théorie de la mesure est la théorie adéquate pour répondre à cette question. L'astuce consiste à dire que, dans un cas comme dans l'autre, on est capable d'assigner à chaque partie de E une masse (correspondant au cacao produit dans cette partie) et que ces assignations vérifient certaines propriétés de compatibilité, listées dans la définition de mesure vue en cours (il n'est pas nécessaire d'avoir cette définition en tête pour lire ce document).

Le paragraphe précédent n'est pas parfaitement exact. Si une partie A de la carte du monde est incluse dans un pays donné, disons le Ghana, alors pour la première modélisation, on voudrait assigner à cette partie la production du Ghana fois l'aire de A divisée par l'aire du Ghana. Or définir l'aire d'une partie A absolument arbitraire du plan n'est pas évident. En fait, on peut démontrer que pour des parties extrêmement pathologiques, il n'y a pas moyen de définir leur aire. Mais nous ne voulons pas abandonner la notion d'aire donc au lieu de chercher à mesurer toutes les parties, on sera content si on arrive à mesurer beaucoup de parties — suffisamment pour qu'aucune des parties qu'on rencontre en pratique ne soit pathologique. C'est pour cette raison qu'on doit introduire les « tribus ».

Dans le reste de ce document, je passerai sous silence les tribus, et ce au motif suivant : les mesures et l'intégration répondent à des questions naturelles, tandis que les tribus répondent à un problème technique interne à la théorie. Les tribus sont donc importantes mais ne sont pas mon propos actuel, qui est de vous dégager une vision des concepts-phares de mesure et d'intégration.

2 Une théorie unificatrice

Les mesures ne sont pas la fin de l'histoire mais le début : quand on a réparti de la masse, ce que permet le choix d'une mesure μ sur un ensemble E , il est alors possible de définir, sous des hypothèses convenables, le nombre $\int f d\mu$ lorsque $f : E \rightarrow \mathbb{R}$. Il m'incombe alors de vous expliquer ce que signifie ce nombre. Le faire en toute généralité n'est peut-être pas le point de départ le plus digeste. Aussi, je préfère lister des mesures μ correspondant à des concepts familiers et indiquer, dans chaque cas, ce à quoi correspond $\int f d\mu$. Bref, entrons en contact avec $\int f d\mu$ par le biais des exemples. De même que je passe les tribus sous silence, je ne donnerai pas les hypothèses à faire sur f car nous n'en sommes pas à ce stade.

Reprenons notre exemple de la section précédente, celui du cacao. On note E la carte du monde. On se donne également μ une mesure encodant la répartition de la production de cacao dans le monde et $f : E \rightarrow \mathbb{R}$ une fonction indiquant, en chaque endroit du globe, le prix du cacao produit à cet endroit, exprimé en €/kg. On peut alors se demander combien cela coûterait d'acheter tout le cacao produit sur Terre (ou dans une zone A donnée). Il semble en effet que la connaissance de μ et de f permette de répondre à cette question, puisqu'on sait quelle est la répartition de la production de cacao ainsi que comment le prix dépend du lieu de production. Eh bien la réponse à cette question est précisément $\int_E f d\mu$ (ou $\int_A f d\mu$ pour une zone A donnée). Quand $A = E$, je note souvent $\int f d\mu$ au lieu de $\int_E f d\mu$.

On pourrait démultiplier les exemples concrets : je ne vais pas emprunter cette voie. Je pense que vous avez déjà en tête qu'une théorie mathématique peut généralement s'appliquer à diverses situations (si ce n'est pas le cas, vous pouvez y méditer). Le point sur lequel je préfère insister relève plutôt du domaine de l'abstrait. Le voici : la théorie de l'intégration telle que présentée dans cette UE comprend comme cas particuliers diverses *théories*. Il s'agit donc d'une théorie très générale et unificatrice.

Le tableau ci-dessous indique, pour divers choix d'ensemble E et de mesure μ , ce que donne la théorie de l'intégration associée : on constate alors qu'on retrouve diverses théories intéressantes comme autant de cas particuliers de la théorie de l'intégration générale étudiée dans cette UE.

Ensemble E	Mesure μ	Fonction $f : E \rightarrow \mathbb{R}$	$\int f d\mu$
\mathbb{N}	Cardinal	$u : \mathbb{N} \rightarrow \mathbb{R}$, notée $(u_n)_{n \in \mathbb{N}}$	Série $\sum_{n=0}^{\infty} u_n$
\mathbb{R}	Longueur	$f : \mathbb{R} \rightarrow \mathbb{R}$	Intégrale simple $\int f(x) dx$
\mathbb{R}^2	Aire	$f : \mathbb{R}^2 \rightarrow \mathbb{R}$	Intégrale double $\iint f(x, y) dx dy$
\mathbb{R}^3	Volume	$f : \mathbb{R}^3 \rightarrow \mathbb{R}$	Intégrale triple $\iiint f(x, y, z) dx dy dz$
Ω	Probabilité	Variable aléatoire $X : \Omega \rightarrow \mathbb{R}$	Espérance $\mathbb{E}[X]$

Pour comprendre la deuxième colonne, il convient d'avoir à l'esprit que quand je disais « répartir de la masse », cette masse n'était pas obligée d'être un poids physique. Cela peut correspondre à n'importe quelle notion raisonnable de « taille » : pour mesurer la taille d'une partie A d'un ensemble E , on peut envisager le cardinal de A si $E = \mathbb{N}$, la longueur de A si $E = \mathbb{R}$, etc. Pour un ensemble E fixé, on peut d'ailleurs envisager diverses notions de taille concurrentes, correspondant à des mesures différentes μ sur ce même ensemble E . C'était le cas dès la première page : si je prends la production de cacao à l'automne 2024 ou au printemps 2025, j'obtiendrai deux mesures différentes sur la carte du monde.

Bref : le saut conceptuel effectué ce semestre sera récompensé par une grande unification, regroupant séries, espérances et intégrales dans un même formalisme.

3 Une théorie puissante

Dans la dernière ligne du tableau vu plus haut, la variable aléatoire peut être discrète ou à densité. Mais elle peut aussi n'être ni discrète ni à densité. Et c'est une des vertus de la théorie générale de l'intégration : elle ne fait pas qu'unifier des choses connues, elle ouvre aussi de nouvelles possibilités.

De plus, même si on ne s'intéresse qu'à l'intégrale $\int_{-\infty}^{+\infty} f(x) dx$, passer au point de vue moderne est avantageux. D'une part, on devient capable d'intégrer beaucoup plus de fonctions : on n'a pas du tout besoin de continuité par morceaux par exemple, des conditions considérablement plus souples suffisent. Et d'autre part, c'est le bon cadre pour démontrer divers théorèmes d'intégration, tels que le théorème de convergence dominée. Quand bien même on souhaiterait ne démontrer ce théorème que pour les fonctions continues par morceaux, rester dans ce cadre serait bien moins commode que de passer par la théorie générale.

La théorie étudiée dans cette UE est donc effectivement très puissante. Elle unifie et complète les théories des séries, intégrales et espérances. Dans le cas où $(u_n)_{n \in \mathbb{N}}$, $f : \mathbb{R} \rightarrow \mathbb{R}$ ou $X : \Omega \rightarrow \mathbb{R}$ sont en fait à valeurs *positives*, la théorie moderne est parfaite et couvre en particulier tout ce que vous avez vu en L1-L2. Quand $(u_n)_{n \in \mathbb{N}}$, f ou X prend des valeurs aussi bien positives que négatives, la théorie moderne ne gère bien les choses que sous une hypothèse dite « d'intégrabilité ». Pour $(u_n)_{n \in \mathbb{N}}$, cela revient à supposer que la série est absolument convergente, c'est-à-dire que $\sum_{n=0}^{\infty} |u_n| < \infty$. Pour $f : \mathbb{R} \rightarrow \mathbb{R}$, cela revient à dire que $\int_{-\infty}^{+\infty} |f(x)| dx < \infty$. Concrètement, il faut penser cette hypothèse comme évitant des problèmes du genre « $\infty - \infty$ ».

Ainsi, il y a tout de même *un* type de questions pour lequel ce cours de L3 ne supprime pas ce que vous avez vu en L2, à savoir les questions sensibles à « l'ordre de parcours ». Je m'explique. On peut trouver $(u_n)_{n \in \mathbb{N}}$ et une bijection $\varphi : \mathbb{N} \rightarrow \mathbb{N}$ telles que les nombres $\sum_{n=0}^{\infty} u_n$ et $\sum_{n=0}^{\infty} u_{\varphi(n)}$ aient un sens du point de vue « cours de série » mais que $\sum_{n=0}^{\infty} u_n \neq \sum_{n=0}^{\infty} u_{\varphi(n)}$. Cela n'arrive jamais si on suppose l'intégrabilité, c'est-à-dire ici la convergence absolue de la série. De même, on peut trouver des fonctions $f : \mathbb{R} \rightarrow \mathbb{R}$ telles que les limites $\int_{-n}^n f(x) dx$ et $\int_{-n}^{n+1} f(x) dx$ existent mais soient deux nombres différents. Cela n'arrive jamais si on suppose que f est intégrable. La théorie générale de l'intégration ne gère pas ces subtilités. Il s'agit d'un prix bien modique à payer pour la pléthore de services rendus par cette théorie. Pour voir les choses de façon pleinement positive : le fait que ces subtilités n'apparaissent pas dans la théorie moderne signifie qu'on travaillera dans un cadre où les choses se passent bien, où ces pathologies ne peuvent jamais venir nous embêter.

4 Qu'est-ce que l'intégrale ?

Dans la section 2, j'ai expliqué à quoi correspondait $\int f d\mu$ pour des choix précis de E et μ . Peut-on se faire une idée de ce qu'il en est lorsque l'ensemble E et la mesure μ sont arbitraires ? Cela est en effet possible, de la façon suivante.

Soit E un ensemble et soit $f : E \rightarrow \mathbb{R}$. Imaginons que je décide de déposer *une* unité de masse sur un élément $x_0 \in E$ et de ne pas mettre de masse ailleurs : il est assez naturel de décréter qu'à cette distribution de masse et à ce choix de fonction f , on décide d'associer le nombre $f(x_0)$. En tout cas, c'est ainsi que je définis mes règles du jeu. Si je décide de déposer *trois* unités de masse sur x_0 et rien ailleurs, dans ce cas, je décrète que le résultat sera $3f(x_0)$. Enfin, si je place *deux* unités de masse sur x_0 et *sept* unités sur x_1 , alors le résultat sera $2f(x_0) + 7f(x_1)$.

Comment continuer cette suite logique ? Si la distribution de masse μ est arbitraire, pas forcément concentrée sur des points mais possiblement bien étalée, comment mimer cette procédure ? Le nombre $\int f d\mu$ est précisément la réponse à cette question.

5 Pourquoi s'intéresser à l'intégrale telle qu'abordée dans ce cours ?

Nous avons vu diverses raisons de s'intéresser à l'intégrale moderne :

- elle fournit une théorie générale des pondérations,
- elle permet un bon traitement de l'intégrale usuelle,
- elle permet un bon traitement de l'espérance.

Je n'inclus pas dans la liste le traitement des séries car, s'il est vrai que les résultats sur les séries peuvent être retrouvés comme cas particuliers à partir de l'intégration générale, on peut les obtenir de façon plus directe sans passer par cette théorie (en outre, ne serait-ce que pour formuler la définition de « mesure », on a déjà besoin d'une compréhension, certes rudimentaire, des séries). Cette réserve ne s'applique pas pour l'espérance et l'intégrale : là, le jeu en vaut vraiment la chandelle. D'ailleurs, en ce qui concerne la théorie des probabilités, cette UE ne sert pas seulement à définir l'espérance : elle offre le socle théorique servant de fondement à toute la théorie moderne des probabilités. De nos jours, pour faire des probabilités de façon approfondie, la théorie de la mesure est incontournable.

Aux raisons précédentes on pourrait en ajouter une dernière, correspondant à des motivations plus algébriques. Intéressons-nous à l'espace vectoriel de toutes les fonctions continues de $[0, 1]$ vers \mathbb{R} . Il s'agit d'un espace vectoriel normé (on considère la norme uniforme). On le notera V , la lettre E étant déjà prise dans ce texte pour désigner d'autres choses. Voici maintenant une légitime question d'algèbre linéaire : quelles sont les formes linéaires continues sur V ? Il se trouve que cette question est très intimement liée à l'étude des mesures sur $[0, 1]$. Nous allons approfondir cela dans la fin de ce document. Cette question est une incarnation technique d'une question plus floue : quelles sont toutes les façons « naturelles » d'associer un nombre réel à chaque élément $f \in V$?

Il est facile de trouver quelques formes linéaires continues sur V . Par exemple, si $x_0 \in [0, 1]$, l'application d'évaluation en x_0 donnée par la formule $f \mapsto f(x_0)$ est une telle forme linéaire. C'est aussi le cas de $f \mapsto 3f(x_0)$ ou encore de $f \mapsto 2f(x_0) + 7f(x_1)$, si x_1 est également un élément de $[0, 1]$.

Mais attendez : ces exemples figuraient déjà en section 4. Et on a vu que l'intégrale cherche précisément à généraliser de façon pertinente ce genre de procédure. Soit donc μ une mesure sur $[0, 1]$. Il est alors tentant de se demander si $f \mapsto \int f d\mu$ définit bien une forme linéaire continue sur V . Il se trouve que c'est le cas si et seulement si la masse totale $\mu([0, 1])$ est un nombre fini, c'est-à-dire ne vaut pas l'infini : on dit qu'une telle mesure est *finie*. Par exemple, en plus des trois exemples précédents, on aurait pu mentionner $f \mapsto \int_0^1 f(x) dx$. Les mesures finies permettent aussi de rendre compte d'exemples plus élaborés, tels que $f \mapsto \int_0^{1/2} f(x) dx + f(1)$ ou encore $f \mapsto \int_0^1 xf(x) dx$.

Bravo : chaque mesure finie sur $[0, 1]$ définit une forme linéaire continue sur V , et dès qu'on se donne deux mesures finies distinctes, cela donne bien lieu à deux formes linéaires distinctes. Cela constitue déjà un lien fort entre l'étude des formes linéaires continues sur V et celle des mesures sur $[0, 1]$. Mais on peut aller plus loin encore. . .

En effet, toutes les formes linéaires Φ ainsi produites sont *positives* : cela signifie que si une fonction $f \in V$ est positive (c'est-à-dire vérifie $\forall x \in [0, 1], f(x) \geq 0$), alors on a $\Phi(f) \geq 0$. Cela revient à une propriété familière de l'intégrale : intégrer une fonction positive donne un résultat positif. Eh bien un théorème dû à Riesz (prononcer "r i s") indique que toute forme linéaire continue sur V qui est positive s'écrit sous la forme $f \mapsto \int f d\mu$ pour une unique mesure finie μ . Combiné avec ce qu'on a dit juste avant, cela signifie qu'on dispose d'une bijection de l'ensemble des mesures finies sur $[0, 1]$ vers l'ensemble des formes linéaires continues sur V qui sont positives, à savoir la bijection qui envoie μ sur la forme linéaire $f \mapsto \int f d\mu$.

Ainsi, une civilisation qui ferait des mathématiques fondamentales sans s'être intéressée à l'intégration pour elle-même ni à l'espérance pourrait tout de même découvrir l'intégration générale, en cherchant à comprendre les formes linéaires continues et positives sur V . Peut-être deviendrait-elle alors curieuse de l'intégration générale et chercherait à étudier les mesures qui ne sont pas finies, à intégrer des fonctions moins régulières que celles appartenant à V , etc. Mais revenons à notre étude.

Pour une forme linéaire Φ , la condition de positivité est équivalente à la condition de croissance, qui dit que si deux éléments f et g de V vérifient $f \geq g$, alors on a forcément $\Phi(f) \geq \Phi(g)$. Comme l'analyse est l'art du petit et du grand, des majorations et des minorations — bref, l'art des inégalités —, il est naturel du point de vue de l'analyse de s'intéresser aux formes linéaires continues sur V qui sont croissantes (ou, ce qui revient au même, positives). Le théorème de Riesz répond donc à des questions finalement assez naturelles du point de vue de l'analyse. Mais si notre civilisation imaginaire était un peu plus algébriste et un peu moins analyste, redécouvrirait-elle la théorie de la mesure ? Bref : si on s'intéresse aux formes linéaires continues sur V , sans imposer la condition de positivité, y a-t-il encore un lien très fort avec les mesures finies sur $[0, 1]$?

La réponse est oui. Si Φ et Ψ désignent deux formes linéaires continues positives sur V , alors $\Phi - \Psi$ est une forme linéaire continue sur V , qui n'est pas nécessairement positive (en fait, cela est vrai même si Φ et Ψ ne sont pas supposées positives, mais je les prends positives car j'ai en tête la suite de mon propos). Ainsi, si μ et ν désignent deux mesures finies sur $[0, 1]$, alors $f \mapsto \int f d\mu - \int f d\nu$ est une forme linéaire continue sur V . Et un théorème indique que toute forme linéaire continue sur V s'écrit en effet sous cette forme.

Toutefois, cette écriture n'est pas unique, et ce pour une raison assez simple à concevoir : si j'ajoute de la masse dans μ et la même masse dans ν , en faisant la soustraction, la masse ainsi introduite ne jouera finalement aucun rôle. Plus précisément, si μ, ν et ρ désignent trois mesures finies sur $[0, 1]$, et si je pose $\tilde{\mu} := \mu + \rho$ et $\tilde{\nu} := \nu + \rho$, alors les formes linéaires $f \mapsto \int f d\tilde{\mu} - \int f d\tilde{\nu}$ et $f \mapsto \int f d\mu - \int f d\nu$ sont égales. Travailler avec $(\tilde{\mu}, \tilde{\nu})$ plutôt que (μ, ν) est possible mais pas idéal : le fait que ρ apparaisse dans $\tilde{\mu}$ et dans $\tilde{\nu}$ introduit d'inutiles redondances. Il se trouve qu'on peut obtenir unicité du couple (μ, ν) si on exige l'absence de redondances — précisons cela au prochain et dernier paragraphe. Il n'y a pas de terminologie standard pour les définitions du prochain paragraphe : j'ai choisi les mots qui me paraissaient appropriés mais ils ne vous seront d'aucun secours pour des requêtes sur un moteur de recherche, et vous devrez les redéfinir si vous souhaitez les employer dans un échange mathématique avec quelqu'un.

Si (μ_1, ν_1) et (μ_2, ν_2) sont deux couples de mesures finies sur $[0, 1]$, disons que (μ_1, ν_1) *se simplifie* en (μ_2, ν_2) s'il existe une mesure finie ρ sur $[0, 1]$ telle que $\mu_1 = \mu_2 + \rho$ et $\nu_1 = \nu_2 + \rho$. Qualifions un couple de *totalelement simplifié* si le seul couple en lequel il peut se simplifier est lui-même. Soit Φ une forme linéaire continue sur V . Disons qu'un couple (μ, ν) de mesures finies sur $[0, 1]$ *représente* Φ si, pour tout $f \in V$, on a $\Phi(f) = \int f d\mu - \int f d\nu$. Il est alors possible de démontrer qu'il existe un unique couple totalelement simplifié représentant Φ . De plus, n'importe quel couple représentant Φ peut se simplifier en ce couple-ci. Ainsi, on dispose d'une bijection de l'ensemble des couples totalelement simplifiés de mesures finies sur $[0, 1]$ vers l'ensemble des formes linéaires continues sur V , à savoir la bijection qui envoie (μ, ν) sur la forme linéaire $f \mapsto \int f d\mu - \int f d\nu$.