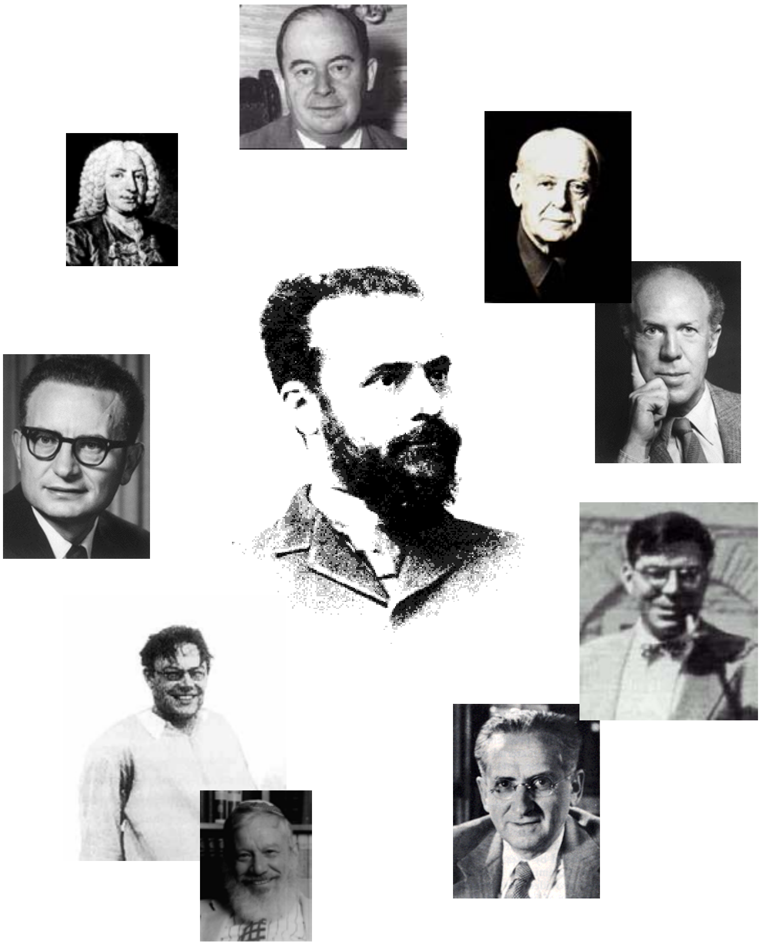
La décision dans l'incertain
préférences, utilité et probabilités

Philippe Bernard

Juillet 2000

Table des matières



Figure~1:

1 Le risque

La prise en compte de la dimension temporelle dans la modélisation permet de réintroduire les comportements financiers, d'épargne et d'investissement ainsi que d'autres fonctions des marchés financiers. Néanmoins, un aspect essentiel de l'activité economique demeure absent: son *risque*.

En effet, une hypothèse centrale des analyses intertemporelles est celle des anticipations exactes. Avec elle, l'évolution future de l'économique est certaine, unique, et tenue comme telle par les agents économiques. Aussi, par exemple, à chaque instant, les revenus engendrés par les actifs financiers, les investissements sont connus. Le risque, l'incertitude sont donc absents de l'analyse; à l'équilibre de ce monde, aucune prime de risque ne peut exister, tous les actifs ont le même rendement net, aucune spéculation ne peut avoir lieu, aucune assurance n'est demandée.

Pour réintroduire l'ensemble des institutions financières, des contrats dont l'objet est de protéger les agents de l'incertitude, il est donc nécessaire de prendre en compte le *hasard*, le *risque*. Mais admettre que les projets sont risqués, c'est admettre qu'*il peut se passer plus de choses qu'il ne s'en passera*. Ainsi, nous nous assurons contre des *événements* rares (incendies, accidents), que nous espérons (et pensons souvent) ne pas voir se produire mais qui *peuvent* arriver. De même, sur le marché des actions, celui de l'immobilier, lorsque nous anticipons une hausse prolongée mais nous n'investissons pas la totalité de notre patrimoine dans ces actifs risqués car nous admettons que ceux-ci *peuvent* également baisser. Bref, comme le suggèrent les racines de hasard et de risque[1], quelles que soient leurs compétences, leurs précautions, les investisseurs, les financiers sont fréquemment dans la situation de joueurs de dés: ils connaissent approximativement les résultats possibles, les fréquences de ceux-ci mais ne sont pas totalement maitres de leurs destins: le *démon de la chance* (M. Kendall) le détermine aussi en partie.

La prise en compte de cette donnée dans l'activité économique en général, les jeux de hasard, l'assurance, la finance en particulier a conduit très tôt les hommes à s'interroger sur la nature de ces aléas, à tenter de les quantifier.

2 Le printemps de l'analyse du risque*

La description des risques économiques, leurs quantifications, i.e. le fait de les résumer par des nombres, a évidemment une histoire aussi longue que le commerce et l'assurance.[2] Cependant, les premières tentatives modernes d'analyse de ce problème remontent à la Renaissance[3] et furent suscitées par des activités plus futiles.

[1] Hasard vient du mot arabe "al zahr" (= dé) et risque vient de l'italien "riscare" oser.

[2] Sur ce sujet, on peut se référer notamment à [Dav62], [Hac75], ou à l'ouvrage plus récent de P.L. Bernstein [Ber96].

[3] Rappelons que l'introduction dans le monde latin est en général attribuée à Leonardo Fibonnaci (1170-1240) dont le traité *Liber Abacci* fut publié en 1202. Fils d'un marchand de Pise, Fibonnaci fut envoyé par celui-ci en Orient pour y apprendre les méthodes de calcul pratiquées en Orient, inconnues des Latins. Il se rendit en Égypte, en Syrie, à Byzance, en Sicile et en Provence. Après son retour à Pise, vers1200, Fibonnaci publia quelques ouvrages, dont le *Liber Abbaci*, destinés à répandre la connaissance qu'il avait acquise du calcul arabo-indien. Ces ouvrages contenaient de nombreuses applications pratiques susceptibles d'intéresser les commerçants: intérêt, profit, change, etc. La diffusion des nouveaux chiffres fut cependant lente et suscita de nombreuses résistances jusqu'au début du XVIe siècle. Ainsi, en 1229, Florence édicta une loi interdisant aux banquiers l'usage des "symboles infidèles".

Figure~2: Gerolamo Cardano

2.1 Probabilités

Le père (relativement) méconnu de la théorie des probabilités est sans doute l'italien Gerolamo Cardono (ou Jérôme Cardan) (1501-1576).

Un des médecins les plus célèbres de son époque, Cadano était surtout un homme de la Renaissance passionné de littérature, de philosophie, de mathématiques[4], d'astrologie, ... et de jeux. Comme il le confessa dans sa biographie *De Vita Propria Liber* [Car30], il conçut "un amour immodéré des jeux de table et de dés [...] Pendant de nombreuses années [...] je ne jouais pas de temps en temps, mais comme je dois l'avouer, chaque jour." ([Car30], cité par [Ber96] p. 45). Au jeu, il préférait cependant s'en remettre plus souvent à la tricherie qu'au hasard et seule la protection du souverain pontife le sauva de la vindicte des joueurs abusés (et des maris trompés). Aussi, dans ses travaux mathématiques sur les jeux de hasard, il prit toujours grand soin de qualifier ses résultats en ajoutant la clause de précaution: "si les dés ne sont pas pipés".

Cardano fut en effet l'un des premiers mathématiciens à analyser les *jeux de hasard* et introduisit les premiers éléments de la théorie des probabilités: "Cardano a sans doute été le premier à introduire la dimension statistique de la théorie des probabilités. [...] Il proposa, pour la première fois, ce qui est aujourd'hui la forme courante pour définir une probabilité comme fraction: le nombre de résultats favorables divisés par le "circuit" - c'est-à-dire le nombre

[4] On lui doit notamment l'introduction des nombres imaginaires.

total de résultats possibles." ([Ber96] p. 49) Il ne parla cependant jamais de "probabilité" mais de "chance". Dans son traité des jeux *Liber de Ludo Alea* (Livre sur les Jeux de Hasard), écrit en 1525, réécrit en 1565, Cardano compila toute une série de fréquence sur les jeux de hasard et reconnut, avant Pascal, l'importance des combinaisons pour la théorie des probabilités. Mais cet ouvrage ne fut publié qu'en 1663.[5]

Un siècle après Cardano, la passion du jeu fit encore progresser la mesure du risque. L'initiateur fut en effet encore cette fois un joueur invétéré: le chevalier de Méré (1610-1685). Le chevalier était un homme heureux en jeu. Mais, à la différence de Cardano, son succès ne devait rien à la malhonnêté. Sans doute instruit par l'expérience, le chevalier semble en effet avoir compris qu'obtenir un six en quatre coups est plus "probable" (=?) que son contraire.[6] En jouant souvent, et donc exploitant sans le savoir la loi des grands nombres, le chevalier de Méré put donc prospérer momentanément. Sa martingale avait cependant un défaut: le gain n'est très probable qu'à long terme: des périodes prolongées de revers sont possibles. Il est nécessaire de pouvoir supporter durablement des pertes; un capital initial important est donc nécessaire. Aussi, tenta-t-il des variantes. Notamment, il s'essaya au *sonnez* (double-six) en 24 lancés. Il perdit alors suffisamment d'argent pour avoir des doutes sur la valeur de cette nouvelle stratégie.[7] Désorienté, le chevalier décida que le temps de l'empirisme était clos, que celui de la réflexion était venu. Aussi décida-t-il de soumettre ses angoisses à une de ses connaissances de jeu: Blaise Pascal (1623-1662).

En 1650, après une première crise mystique, Pascal était tombé gravement malade: provisoirement partiellement paralysé, il souffrait également de maux de tête que ses médecins étaient incapables de soigner. En désespoir de cause, ils lui conseillèrent de renoncer à sa nouvelle vie d'ascète et de

[5] Galilée (1554-1642) a aussi consacré un petit traité aux jeux de hasard *Sopra le Scoperte dei Daddi* (Jouer aux dés), publié en 1623, dédié à son protecteur Cosme II, le Grand Duc de Toscane, un joueur passionné. Comme Cardano, Galilé s'occuppa essentiellement de la fréquence des événements, des différents tirages des jeux de hasard. Apparemment cependant en 1623 cette conception des probabilités était déjà courante car Galilé ne prétendit pas faire oeuvre originale.

[6] En effet, comme la probabilité de ne pas obtenir 6 à un lancé est $5/6$, la probabilité de n'obtenir aucun 6 en quatre lancés est $(5/6)^4 = 625/1296$. Le complémentaire de cet événement, le fait d'obtenir au moins un six, est donc $1 - \frac{625}{1296} = 0.51775$.

[7] En effet, la probabilité d'obtenir un double six étant $1/36$ en un lancé, la probabilité de n'obtenir aucun double six est $35/36$ en un lancé, $(35/36)^{24}$ en 24 lancés. Par conséquent, la probabilité d'obtenir un double six en 24 lancés n'est que de 0.4914. Si le chevalier avait eu l'heureuse idée de parier en 25 lancés, les probabilités auraient été cette fois en sa faveur. En effet, par les mêmes calculs, la probabilité d'obtenir un double-six en 25 lancés est 0. 50553. A un lancé près, l'histoire eut peut-être alors été différente...

Figure~3: Blaise Pascal et Pierre de Fermat

reprendre sa vie antérieure de libertin; Pascal obtempèra, renoua avec ses anciennes amitiés de jeux, en noua de nouvelles. Il rencontra alors le chevalier de Méré qui lui soumit ses problèmes: la recherche d'un six en moins de 4 lancés est-elle une stratégie meilleure que la recherche du double-six en moins de 24 lancés?

Le problème parut si formidable à Pascal qu'il hésita d'abord, puis décida de s'associer à Pierre de Fermat (1601-1665) pour résoudre ce problème. Ensemble, dans leur correspondance, ils jetèrent en 1654 les fondements de la théorie des probabilités, voire ceux du *risk management*. La contribution de Pascal fut notamment de redécouvrir (après Cardano) l'importance des combinaisons pour la théorie des probabilités et de proposer, pour le calcul des combinaisons, le triangle qui porte son nom.[8]

Selon l'historien des sciences Ian Hacking [Hac75], la seconde contribution majeure de Pascal fut de fonder dans son fragment 418 de ses *Pensées* (sur le pari) la théorie de la décision, i.e. "la théorie de la décision lorsque le futur est incertain".

2.2 Le pari de Pascal comme problème de décision

Dans ce texte, Pascal pose initialement le principe de l'impossibilité de prouver ou l'inexistence de Dieu:

[8]Celui-ci en fait avait été formulé pour la première fois en 1303 par un mathématicien chinois, Chu Shih-chieh, mais était resté inconnu des occidentaux.

> "S'il y a un Dieu il est infiniment incompréhensible, puisque n'ayant ni parties ni bornes, il n'a nul rapport à nous. Nous sommes donc incapables de connaître ni ce qu'il est, ni s'il est. Cela étant qui osera entreprendre de résoudre cette question? Ce n'est pas nous qui n'avons aucun rapport à lui." ([Pas62] pp. 175-6)

Mais, dans ce problème, l'homme est comme un joueur engagé dans un jeu de hasard:

> "Dieu est ou il n'est pas; mais de quel côté pencherons-nous? la raison n'y peut rien déterminer. Il y a un chaos infini qui nous sépare. Il se joue un jeu à l'extrémité de cette distance infinie, où il arrivera croix ou pile." ([Pas62] pp. 176)

Même s'il lui est impossible de répondre à la question de l'existence de Dieu, l'homme est contraint par sa condition de choisir:

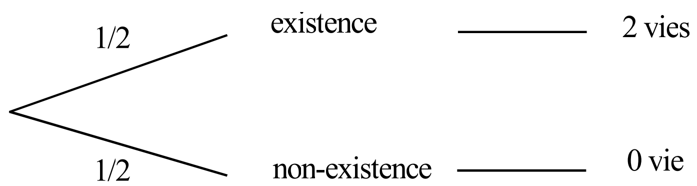
> "Oui, mais il faut parier. Cela n'est pas volontaire, vous êtes embarqués. Lequel prendrez-vous donc? " ([Pas62] p. 176)

Pour choisir, Pascal propose de peser les avantages et les inconvénients de chaque option comme tout parieur:

> "Voyons; puisqu'il faut choisir voyons ce qui vous intéresse le moins. [...] Tout joueur hasarde avec certitude pour gagner avec incertitude, et néanmoins il hasarde certainement le fini pour gagner incertainement le fini, sans pécher contre la raison. [...] l'incertitude de gagner est proportionnée à la certitude de ce qu'on hasarde selon la proportion des hasards de gain et de perte. Et de là vient s'il y a autant de hasards d'un côté que de l'autre le parti est à jouer égal contre égal." ([Pas62] pp. 176-78)

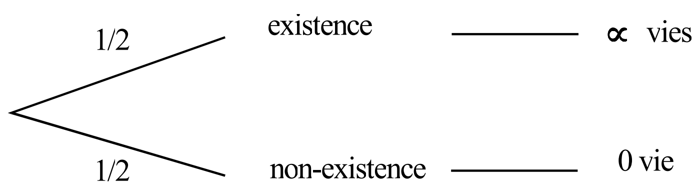
Pascal admet que parier l'inexistence de Dieu puisse être une stratégie optimale pour certaines loteries notamment pour celle représentée sur la figure 4:

> "Voyons puisqu'il y a pareil hasard de gain et de perte, si vous n'aviez qu'à gagner deux vies pour une, vous pourriez encore gager[.]" ([Pas62] p. 176)



valeur de la loterie = 0.5 x 2 + 0.5 x 0 = 1 vie

Figure~4: La première loterie analysée par Pascal



valeur de la loterie = 0.5 x ∝ + 0.5 x 0 = ∝ vies

Figure~5: Le jeu pertinent pour Pascal

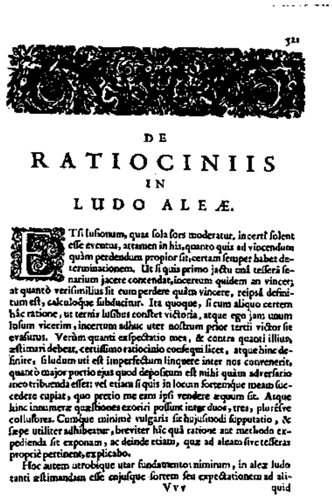
Mais, la contrepartie de la croyance en Dieu n'est pas une simple vie supplémentaire mais l'éternité. Aussi, le contexte du problème est non celui de la figure 4 mais celui de la figure 5, et donc le problème se pose dans les termes suivants:

> "il y a ici une infinité de vie infiniment heureuse à gagner, un hasard de gain contre un nombre fini de hasards de perte et ce que vous jouez est fini. [...] Il n'y a point à balancer, il faut tout donner. Et ainsi quand on est forcé à jouer, il faut renoncer à la raison pour garder la vie plutôt que de la hasarder pour le gain infini aussi prêt à arriver que la perte du néant." ([Pas62] p. 177)

Et Pascal de conclure quelque peu cyniquement:

> "Vous voulez aller à la foi et vous n'en savez pas le chemin. Vous voulez vous guérir de l'infidélité et vous en demandez les remèdes, apprenez de ceux, etc. qui ont été liés comme vous et qui parient maintenant tout leur bien. Ce sont gens qui savent ce chemin que vous voudriez suivre et guéris d'un mal dont vous voulez guérir; suivez la manière par où ils ont commencé. C'est en faisant tout comme s'ils croyaient, en prenant de l'eau bénite, en faisant dire des messes, etc. Naturellement même cela vous fera croire et vous abêtira." ([Pas62] p. 178)

L'oeuvre de Pascal et de Fermat influença rapidement de nombreux contemporains: *Christiaan Huygens* publie en 1657 son traité sur les probabilités *De Ratiociniis in Ludo Aleae* (Sur le raisonnement dans les jeux de dés), *Leibniz* applique le nouvel outil aux problèmes de justice, les membres de Port Royal publient *Ars Cogitandi* (La logique ou l'art de penser) dont l'auteur principal est sans doute *Antoine Arnauld*. La dernière partie de cet ouvrage comprend quatre chapitres sur les probabilités exposant notamment le principe de l'inférence statistique. Analysant les comportements, les auteurs relevèrent que si la probabilité d'être frappée par la foudre est faible, "beaucoup de personnes [...] sont excessivement effrayés quand ils entendent le tonnerre" et en concluent que "la peur d'être frappé doit être proportionnelle non seulement à la gravité du mal, mais aussi à la probabilité de l'événement." Une idée simple mais importante était ainsi mise à jour: la décision est conjointement déterminée par ses résultats possibles et par leurs probabilités.

321

DE

RATIOCINIIS

IN

LUDO ALEÆ.

ETsi lusionum, quas sola fors moderatur, incerti solent esse eventus, attamen in his, quanto quis ad vincendum quàm perdendum propior sit, certam semper habet determinationem. Ut si quis primo jactu unâ tesserâ senarium jacere contendat, incertum quidem an vincet; at quantò verisimilius sit eum perdere quàm vincere, reipsâ definitum est, calculoque subducitur. Ita quoque, si cum aliquo certem hâc ratione, ut ternis lusibus constet victoria, atque ego jam unum lusum vicerim, incertum adhuc uter nostrum prior tertii victor sit evasurus. Verùm quanti exspectatio mea, & contra quanti illius, æstimari debeat, certissimo ratiocinio consequi licet, atque hinc definire, si ludum uti est imperfectum linquere inter nos convenerit, quantò major portio ejus quod depositum est mihi quàm adversario meo tribuenda esset: vel etiam si quis in locum sortemque meam succedere cupiat, quo pretio me eam ipsi vendere æquum sit. Atque hinc innumeræ quæstiones exoriri possunt inter duos, tres, pluresve collusores. Cumque minimè vulgaris sit hujusmodi supputatio, & sæpe utiliter adhibeatur, breviter hîc quâ ratione aut methodo expedienda sit exponam, ac deinde etiam, quæ ad aleam sive tesseras propriè pertinent, explicabo.

Hoc autem utrobique utar fundamento: nimirum, in aleæ ludo tanti æstimandam esse cujusque sortem seu expectationem ad ali-

Vvv quid

Figure~6: Christiaan Huygens et la première page de son ouvrage "De Ratiociniis in Ludo Aleae"

2.3 L'espérance morale

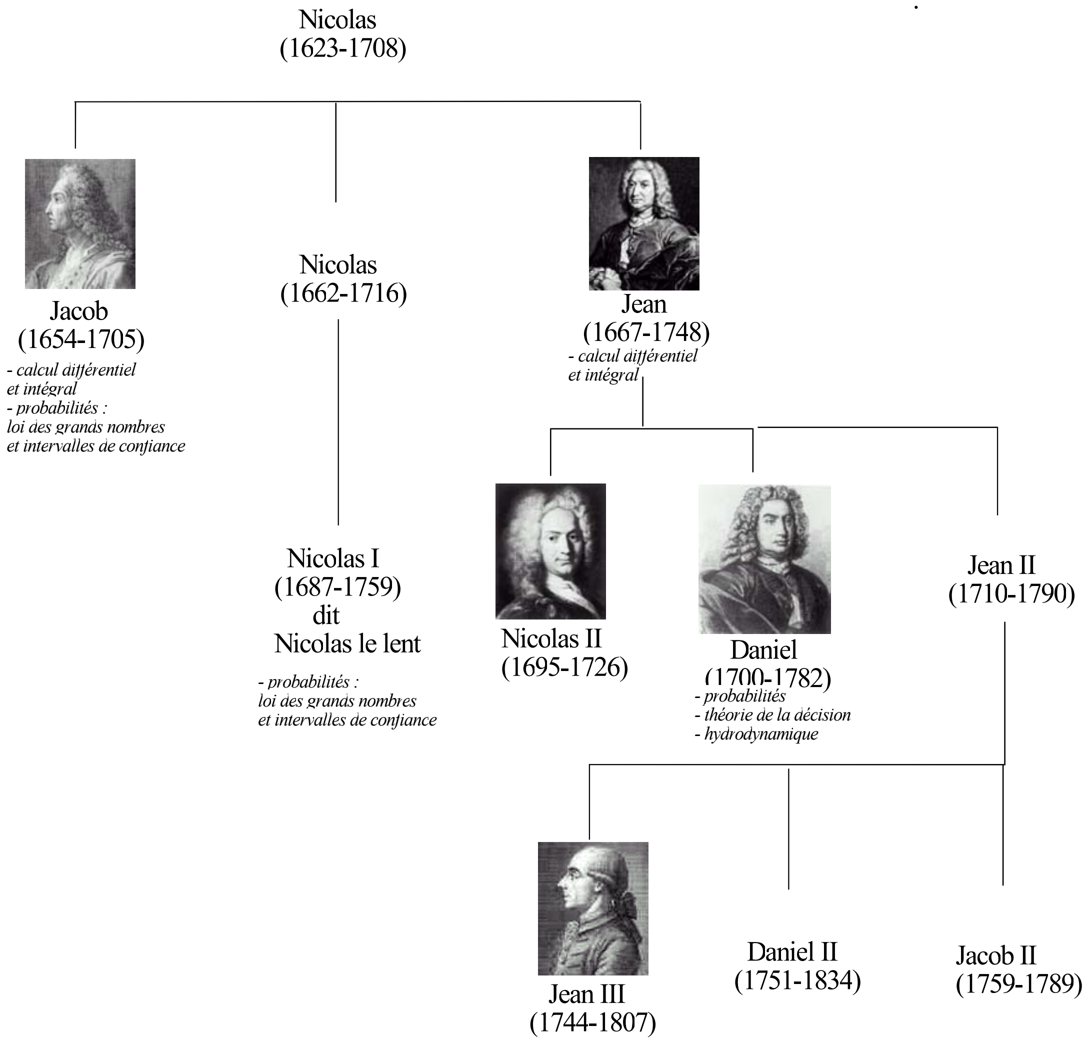
Nicolas Bernoulli (1687-1759), charitablement surnommé par ses contemporains Nicolas le lent, était le membre d'une grande famille de mathématiciens suisses, une des plus grandes dynasties de l'histoire des mathématiques. En 1713, il soumis à la communauté scientifique un problème qui fut à l'origine d'une avancée essentielle de la théorie de la décision dans l'incertain.

Le problème soumis par Nicolas le lent est le suivant. Pierre propose à Paul un jeu à pile ou face (représenté sur la figure 8):

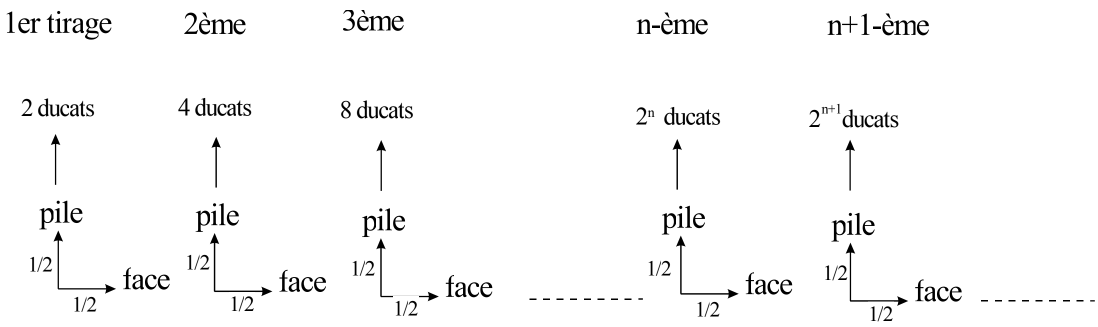
1. si pile arrive au premier lancé, Pierre donnera à Paul 2 ducats[9], sinon le jeu continuera;
2. si pile arrive au second lancé, Pierre donnera à Paul 4 ducats, sinon le jeu continuera;
3. si pile arrive au $n^{ème}$ coup, Pierre donnera à Paul 2^n ducats, sinon le jeu continuera.

Combien vaut donc pour Paul la jeu proposée par Pierre, i.e. quel est le prix maximum qu'il serait prêt à payer pour avoir le droit de participer au jeu?

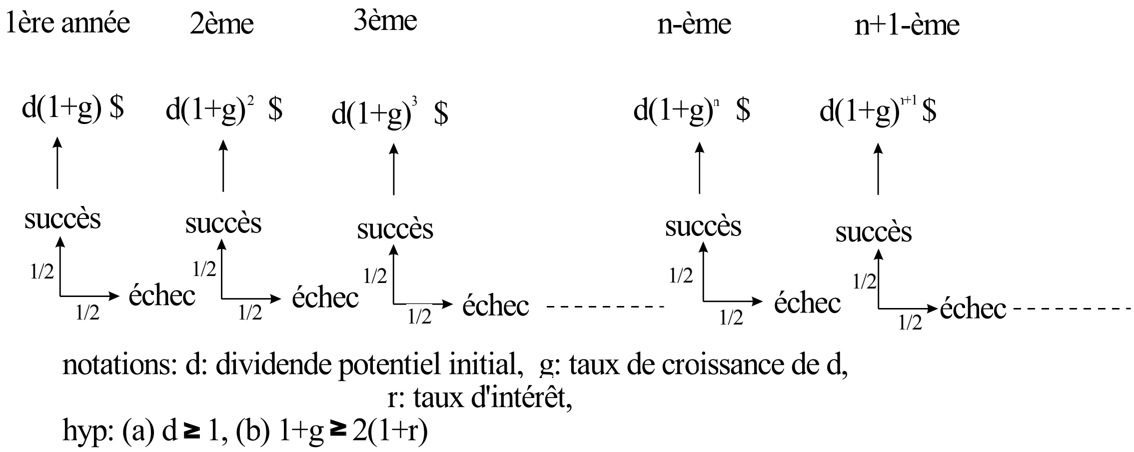
[9] D'après les calculs de P.-L. Bernstein [Ber96], un ducat du XVIIIe siècle est environ l'équivalent de 40$ d'aujourd'hui, 260 FF.



Figure~7: La dynastie des Bernoulli



Figure~8: Le jeu du paradoxe de St Petersbourg

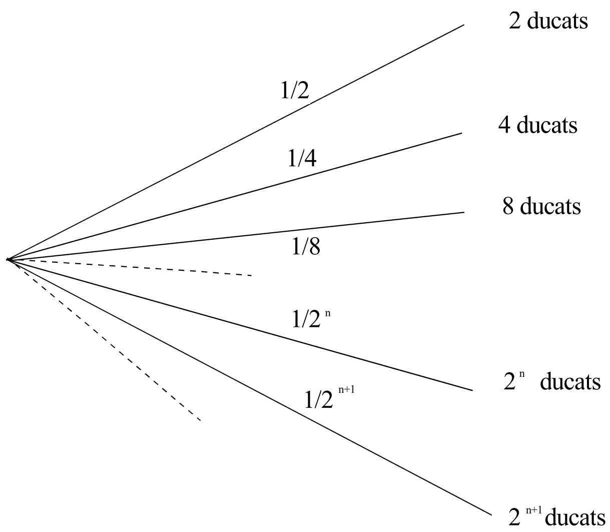


Figure~9: Le problème des actions de croissance

Comme le souligna en 1957 un professeur de finance du MIT, *David Durand*, dans le Journal of Finance, ce jeu n'est pas qu'une curiosité intellectuelle. Il représente un problème financier extrêmement important: l'évaluation des "actions de croissance". Que valent les actions d'entreprise dont on sait à l'avance qu'elles peuvent être très rentables, car elles appartiennent à des secteurs porteurs, mais qui demeurent risquées car elles ne figureront pas forcément parmi les élus?

La figure 9 représente une action dont le dividende est susceptible de connaître une croissance très forte pendante une période de temps substantielle (en fait infini). Le titre demeure risqué: à chaque période, on a une chance sur 2 que l'activité ne dégage aucun profit. Que vaut cette action?

Au XVIIIème, le critère courant utilisé pour évaluer les jeux de hasard était l'espérance du gain. Le jeu proposé par Nicolas le lent peut être repré-



Figure~10: Représentation équivalente du jeu du paradoxe de St Petersbourg

senté sous la forme équivalente de la figure 10. Par conséquent, l'espérance du gain, notée V, est la suivante:

$$V = \frac{1}{2} \times 2 + \frac{1}{4} \times 4 + \frac{1}{8} \times 8 + ... + \frac{1}{2^n} \times 2^n + \frac{1}{2^{n+1}} \times 2^{n+1} + ... = \infty$$

L'application du critère d'espérance du gain débouche donc sur le '*paradoxe de Saint Péterbourg*': quel que soit le prix du jeu fixé par Pierre, Paul selon la théorie de Pascal devrait l'accepter puisque la valeur de celui-ci est infinie. De même, pour la valeur de croissance de la figure 9, si l'on applique *de manière ad hoc* la méthode d'actualisation au gain espéré, la valeur de l'action est:

$$\begin{aligned} V &= \frac{1}{2}\frac{d(1+g)}{1+r} + \frac{1}{4}\frac{d(1+g)^2}{(1+r)^2} + \frac{1}{8}\frac{d(1+g)^3}{(1+r)^3} + ... + \frac{1}{2^n}\frac{d(1+g)^n}{(1+r)^n} + ... \\ &= d.\sum_{j=1}^{\infty}\left(\frac{1}{2}\frac{1+g}{1+r}\right)^j \\ &\geq d.(1+1+...) \end{aligned}$$

Or comme, par hypothèse, $1 + g \geq 2(1 + r)$,

$$\sum_{j=1}^{\infty}\left(\frac{1}{2}\frac{1+g}{1+r}\right)^j \geq 1 + 1 + ... = \infty$$

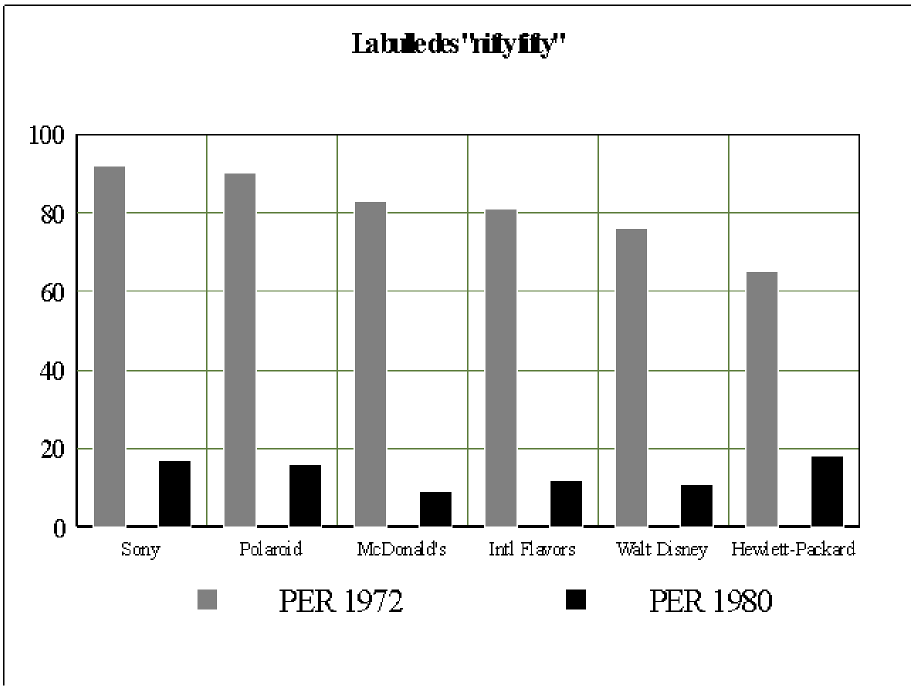
La valeur de l'action de croissance serait alors également indéfinie. Un investisseur adoptant ce critère serait donc prêt à payer n'importe quel prix pour cette action, à trouver normaux des PER extraordinaires, à être plus sensible au risque de "ne pas en être" qu'à celui de "ne pas en avoir assez pour son argent". Pour un praticien comme P.L. Bernstein [Ber95], ceci résume bien certains épisodes boursiers dont le célèbre exemple des "nifty fifty" ("50 favorites")[10].

A la fin des années 60 et au début des années 70, le marché connut un profond engouement pour certaines valeurs de croissance (les "50 favorites") dont les stars étaient des compagnies comme Xerox, Coca-Cola, IBM, Polaroid. Inversement, d'autres valeurs de père de famille comme Union Carbide, General Motors étaient sous-cotées. Aussi, en décembre 1972, même si le dividende moyen versé par les "50 favorites" représentait moins de 50% de la moyenne des valeurs du Standard & Poor's Index 500, le PER de Polaroid était 96, celui de McDonald 80, celui de IFF 73 alors que celui du Standard & Poor's Index 500 était seulement de 19. Le raisonnement des investisseurs semble avoir été souvent le suivant: "quelle importance si l'on payait un prix temporairement trop élevé? Puisque ces actions étaient des actions de croissance, tôt ou tard le prix payé se révélerait justifié." ([Mal96] p.75) En effet, "les gérants de portefeuille de Wall Street n'accordèrent aucune importance mise en garde [de Durand]. Leur histoire d'amour avec les actions de croissance déclencha le boom des "50 favorites" des années 60. Les gérants, qui pensaient agir prudemment, poussaient les prix de ces actions toujours plus haut en les évaluant sur la base d'extrapolations à l'infini de la croissance récente de leurs bénéfices qui étaient extrêmement naïves. Cette pratique stupide coûta des milliars de dollars aux investisseurs, tel que ma fondation caritative, lorsque la réalité s'imposa enfin." ([Ber95] p. 158) En effet, comme le montre la figure 11, la crise boursière de 1974 et les années qui suivirent virent un réajustement brutal de l'ordre des choses.

Le jeu proposé par Nicolas le lent semble donc remettre en cause la pertinence du critère d'espérance des gains. Peu de personnes seraient prêt à payer une somme infinie. Mais quel critère peut remplacer l'espérance du gain? Ce problème posé en 1713 resta (publiquement) sans solution jusqu'à la publication en 1738 d'un mémoire *Specimen Theoriae Novae de Mensura Sortis* (Exposé d'une nouvelle théorie du risque) présenté à l'Académie des Sciences de Saint Pétersbourg.

L'ironie voulut que l'auteur en soit le propre cousin de Nicolas le lent,

[10] Mais on pourrait aussi parler de la bulle sur les valeurs électroniques au début des années 60, le "tronic boom", ou de celui des valeurs biotechnologiques au début des années 80. Les symptômes, le déroulement et l'épilogue sont identiques.



Figure~11: Le réajustement des PER lors de la crise des années 70.

Daniel Bernoulli (1700-1782). Son objectif était non seulement de résoudre le problème de son cousin, mais aussi établir des “règles [qui] peuvent être utilisée par toute personne devant estimer toute prise de risque dans des circonstances financières spécifiques.” [11] Bref, Bernoulli peut être regarder comme le fondateur du “risk management.” Ce faisant il est aussi un des premiers responsables de l’introduction dans la balbutiante théorie de la décision d’un monstre encore mal défini mais incontournable: l’utilité.

Dès le début son mémoire, Bernoulli rappelle la procédure courante pour évaluer:

> “*La valeur espérée est calculée en multipliant chaque gain possible par le nombre de fois où il se produit, et en divisant la somme de ces produits par le nombre total de cas possibles lorsque chaque cas a la même probabilité.*” ([Ber68] p. 15)

et souligne une caractéristique de cette procédure:

> “Aucune caractéristique des personnes n’est prise en considération; seuls comptent les termes du pari[.]” ([Ber68] p. 15)

[11] Cet extrait ainsi que ceux qui suivent sont des traductions de [Ber68].

Figure~12: Daniel Bernoulli

Or, pour lui:

> "la détermination de la *valeur* d'un objet ne doit pas être basée sur ses *avantages*[12], mais seulement sur l'*utilité* qu'il procure. Les avantages de l'objet dépend seulement de lui-même et sont les mêmes pour tout le monde; l'utilité, par contre, dépend des caractéristiques propres de la personne qui fait l'évaluation. Ainsi, il n'y aucun doute qu'un gain de 1000 ducats est sans doute plus apprécié par un pauvre que par un homme riche même si le gain est le même pour les deux." ([Ber68] p. 16)

Même s'il admet que la richesse n'est pas le seul paramètre important de l'utilité[13], il avance son hypothèse fondamentale:

> "[L'] *utilité resultant de tout petit accroissement de la richesse sera inversement proportionnel à la quantité de biens antérieurement possédés.*"

[12] Bernoulli parle en fait ici de "prices" qui correspond dans le cas d'une loterie aux gains monétaires de celle-ci.

[13] "l'utilité d'un objet peut changer avec les circonstances. Ainsi, même si un gain donné est généralement plus utile à un pauvre qu'à un riche, il est néanmoins concevable, par exemple, qu'un riche prisonnier possédant 2000 ducats mais ayant besoin de 2000 ducats supplémentaires pour racheter sa liberté, accordera alors plus de valeur à un gain inopiné de 2000 ducats qu'un homme plus pauvre que lui."([Ber68] p. 16)

Autrement-dit, pour tout accroissement faible de la richesse Δw, l'accroissement de l'utilité (Δu) est alors donné par:

$$\Delta u \approx \frac{1}{w} \times \Delta w \tag{1}$$

Pour Bernoulli, cette hypothèse est le plus souvent valide: "Considérant la nature de l'homme, il me semble que cette hypothèse est valide pour la majorité des hommes auquels ces considérations pourraient être appliquées." ([Ber68] p. 16) Mais cette restriction détermine la fonction d'utilité:

$$\Delta u \approx \frac{1}{w} \times \Delta w \Rightarrow \frac{\partial u}{\partial w} = \frac{1}{w} \Rightarrow u = \ln(w)$$

Ceci suggère la fonction objectif à substituer à l'espérance du gain:

> "*Si l'utilité de chaque profit possible est multiplié par le nombre de fois où il peut être obtenu, et si nous divisons la somme de ces produits par le nombre total de cas possibles, une espérance morale sera obtenue, et le profit qui lui correspond sera la valeur du risque en question.*" ([Ber68] p. 16)

A la maximisation de l'espérance mathématique Bernouilli substitue la maximisation de l' '*espérance morale*': les connaissances sont données par des probabilités $(\pi_i)_{i=1,..}$ - avec $0 \leq \pi_i \leq 1$, $\sum_i \pi_i = 1$ - mais les goûts sont résumés par une fonction d'utilité de la richesse totale (u) quant elle est certaine; en présence d'incertitude, l'agent maximise l' 'espérance morale' U:

$$U = \sum_i \pi_i . u(w + w_i) \tag{2}$$

où w est la richesse certaine initiale, (w_i) les revenus risqués possibles. Ce faisant:

> "[e]n moins d'une page, Bernoulli était passé de l'introduction des probabilités dans les décisions à la prise en compte d'éléments subjectifs dans les décisions en environnements aléatoires [...] Pour la première fois de l'histoire, Bernoulli appliquait la mesure à quelque chose qui *ne pouvait être mesurée.* Il agissait comme un intermédiaire unissant l'intuition et la mesure. Cardano, Pascal, et Ferma avaient donné une méthode pour représenter les aléas de chaque lancé de dé, mais Bernoulli introduisait le *preneur* de risque - le joueur qui agit. Ceci était un domaine d'étude entièrement nouveau." ([Ber96] p. 105)

Comme Bernouilli le remarqua dans les deux corollaires de son mémoire, cette nouvelle approche ne rejette pas l'espérance du gain: elle l'englobe. En effet, si u_i est proportionnelle à la richesse:

$$u(w_i) = \theta.w + \theta.w_i$$

avec θ une constante strictement positive, alors l'espérance du gain demeure le critère (corollaire 1 de Bernoulli):

$$\begin{aligned} U &= \theta w + \theta \sum_i \pi_i w_i \\ &= \theta w + \theta \mathbf{E}(\widetilde{w}) \end{aligned}$$

où $\mathbf{E}(\widetilde{w})$ est l'espérance de la richesse. Enfin, si le risque est suffisamment faible, alors l'espérance demeure une bonne approximation (corollaire 2):

$$u(w + w_i) \approx u(w) + u'(w) \times w_i$$

et donc:

$$\begin{aligned} U &= \sum_i \pi_i u(w + w_i) \\ &\approx \sum_i \pi_i [u(w) + u'(w) \times w_i] \\ &= u(w) + u'(w) \times \sum_i \pi_i.w_i \\ &= u(w) + u'(w) \times \mathbf{E}(\widetilde{w}) \end{aligned}$$

Dans ce cas, chercher le plus grand possible U est équivalent à maximiser $\mathbf{E}(\widetilde{w})$.

Une des premières application de la nouvelle approche est évidemment le paradoxe de Saint Pétersbourg. Dans ce jeu, le "gain" de Paul, évaluée avec l'"espérance morale" n'est plus indéfinie[14] puisque, en supposant une richesse intiale nulle ($w = 0$):

$$\begin{aligned} U &= \lim_{T\to+\infty} \textstyle\sum_{t=1}^{T} \dfrac{1}{2^t} \ln\left(2^t\right) \\ &= \ln(2) \times \lim_{T\to+\infty} \textstyle\sum_{t=1}^{t=T} \dfrac{t}{2^t} \\ &\approx \ln(4) \end{aligned}$$

[14] Bernoulli dans son mémoire porta à la connaissance du public que 10 ans avant lui le mathématicien Gabriel Cramer (1704-1752) avait dans une lettre à Nicolas Bernoulli apporté des solutions très proches des siennes. Comme lui, Cramer mettait en avant l'utilité. Deux solutions différentes étaient proposées pour résoudre le paradoxe: (a) l'utilité est croissante de la richesse, elle peut lui être proportionnelle mais elle est bornée; (b) l'utilité est la racine carrée. Avec cette restriction, la valeur du jeu est en effet finie.

La valeur du jeu proposé par Paul est en fait exactement $\ln(4)$ ducats.

L'étude de son espérance morale conduisit Bernoulli à relever certaines propriétés remarquables. Ainsi, à la différence de l'espérance de gain:

> "il apparaît que dans de nombreux jeux, même ceux qui sont parfaitement équitables, tous les joueurs préféreront subir une perte [plutôt que d'y participer]; ceci constitue constitue l'aversion de la Nature à l'égard du hasard... Ceci est la conséquence de la concavit."([Ber68] p. 20)

Ce résultat est représenté sur la figure 13 pour une loterie équiprobable prenant deux valeurs. Dès lors que la fonction u est concave, tout individu possédant une richesse certaine w et acceptant une telle loterie voit son revenu prendre deux valeurs possibles de part et d'autre de w. S"il perd, l'utilité de l'agent est donnée par l'ordonnée du point B. S'il gagne, elle est donnée cette fois par l'ordonnée du point H. Puisque les probabilités du gain et de la perte sont $1/2$, l'espérance morale est alors la moyenne; graphiquement, elle est donc l'ordonnée du point P, le milieu de la corde BH. Alors que l'espérance de gain de la loterie est donc nulle, l'individu perd donc ici à accepter la loterie puisque son "espérance morale" passe de $u(w)$ à U sur le graphique. Comme Bernoulli le remarqua, l'individu est même prêt à accepter une perte: pour lui, mieux vaut en effet à accepter une baisse de son revenu certain tant que celui-ci dépasse le niveau w' (représenté sur la figure). Cette répulsion à supporter des risques définit une demande d'assurance.

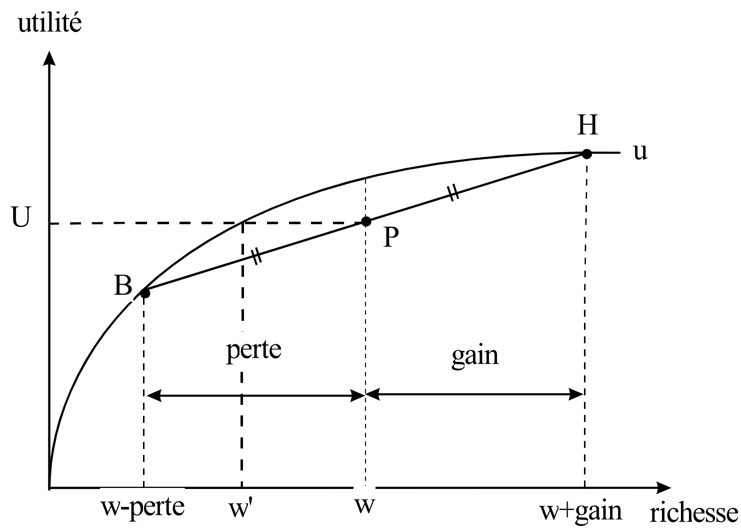
Dans son mémoire, l'étude de celle-ci est la dernière contribution de Bernouilli. Celui-ci montre notamment comment, en présence d'une inégalité des richesses, un marché du risque peut accroître l'espérance morale de chacun des participants.

L'exemple de Bernouilli est le suivant: Caius (?!) réalise un investissement de 10,000 roubles en achetant des marchandises à Amsterdam pour les vendre à St. Pétersbourg. Malheureusement, 5% des expéditions coulent régulièrement dans la Baltique. Si l'on note w_N le revenu de Caius en cas de naufrage, w_V son revenu dans l'autre cas (et donc après la vente à St. Pétersbourg), son espérance morale est donc:

$$U = \frac{5}{100}\ln(w_N) + \frac{95}{100}\ln(w_V)$$

Si l'on note x, la fortune de Caius hors le revenu aléatoire de l'investissement, l'espérance morale de Caius s'écrit aussi:

$$\frac{5}{100}\ln(x) + \frac{95}{100}\ln(x + 10,000)$$



Figure~13: L'inéquité des loteries équitables

En supposant que le coût de l'assurance est de 800, s'il s'assure, il obtiendra alors avec certitude le revenu $x + 9200$. Par conséquent, tant que le prix de l'assurance est de 800, l'individu ne sera demandeur d'une assurance que si:

$$\frac{5}{100}\ln(x) + \frac{95}{100}\ln(x + 10,000) \leq \ln(x + 9200)$$

ou encore en raison des propriétés du ln:

$$\ln\left[x^{\frac{5}{100}}(x + 10,000)^{\frac{95}{100}}\right] \leq \ln(x + 9200)$$

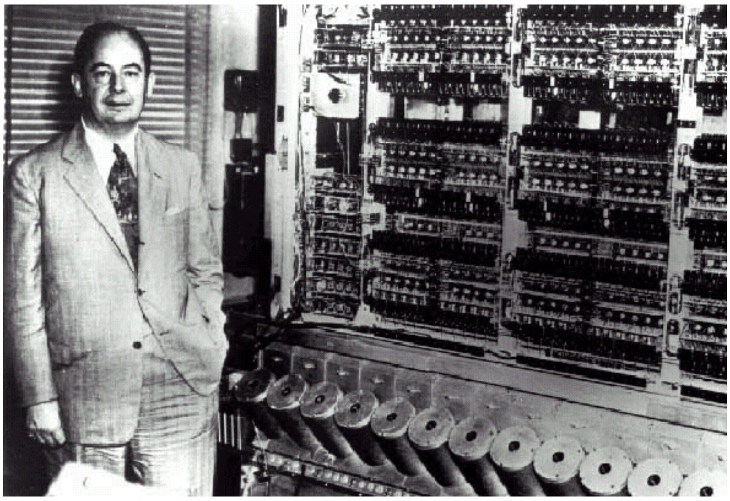
$$x^{5}(x + 10,000)^{\frac{95}{100}} \leq x + 9200$$

i.e. que:

$$x \leq 5043$$

Caius demandera une assurance lorsque la part de son revenu aléatoire représentera une part suffisamment importante de sa fortune: "Nous devons souligner cette vérité, bien qu'elle soit évidente: l'imprudence d'un joueur est d'autant plus grande que la part soumise au hasard de sa fortune est importante."([Ber68] pp. 20-21) L'auto-assurance n'est la stratégie optimale que pour des individus suffisamment riches.

Côté offre, si le prix de l'assurance est de 800, un agent offrant d'assurer Caius aura un revenu égal en l'absence de naufrage à $800 + y$, où y est son revenu initial, à $800 + y - 10,000$ en cas de naufrage puisqu'alors il devra



Figure~14: John von Neumann, génie scientifique polyvalent, fut un des pères de l'informatique, de la bombe à hydrogène ... et de l'économie mathématique

indemniser totalement Caius pour la perte de sa cargaison. Par conséquent, offrir une assurance à Caius ne sera avantageux que si:

$$\frac{5}{100}\ln(y+9,200)+\frac{95}{100}\ln(y+800)\geq\ln(y)$$

ou encore

$$(y-9200)^{\frac{5}{100}}(y+800)^{\frac{95}{100}}\geq y$$

i.e. que:

$$y\geq 20,478$$

Tant que le prix de l'assurance est 800, la mise en place d'un marché de l'assurance sera bénéfique à la fois aux agents très pauvres, dont la richesse, est inférieure à 5400, qui pourront se décharger de leurs risques ainsi qu'aux agents très riches, dont les revenus sont supérieurs à 20478.

De la correspondance de Pascal et de Fermat à Daniel Bernoulli, à peine 80 ans se sont écoulés. La théorie des probabilités ainsi que les statistiques ne vont cesser de progresser. Cependant, la théorie de la décision va tomber dans une certaine torpeur jusqu'au renouveau de la littérature de la théorie des jeux, au début du XXe siècle, avec Emile Borel, John von Neuman notamment. Comme pour de nombreux autres pans de la science économique, le renouveau de la théorie de la décision sera accéléré par la parution en 1944 de *Theory of Games and Economic Behaviour* de J. von Neumann et O. Morgenstern où une axiomatique complète (et ordinale) de la théorie bernouillienne fut pour la première fois donnée.

Figure~15: Vilfredo Pareto et Henri Poincaré

3 Préférences et utilités

3.1 L'approche parétienne

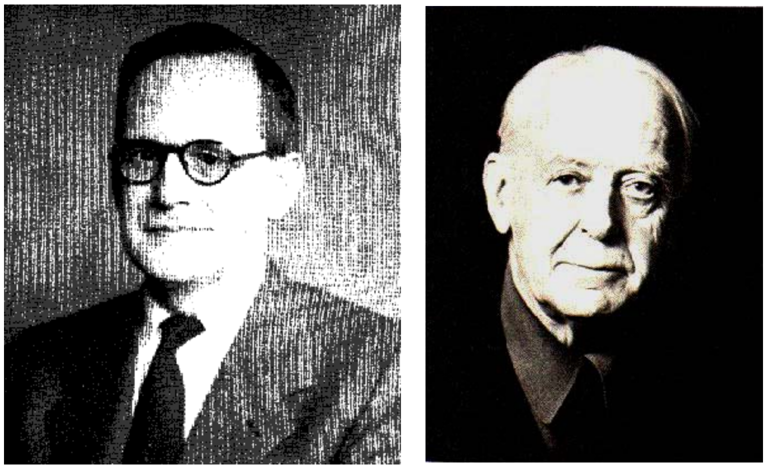
Avant d'aborder les choix dans l'incertain, rappelons quelques éléments de la théorie traditionnelle des choix. Si l'on suppose qu'il existe L biens, l'espace des biens est $\Re^L$. Le consommateur considéré peut choisir ses consommations dans son *ensemble de consommation*, noté X, un sous-ensemble de $\Re^L_+$ résumant à la fois les contraintes propres aux biens (indivisibilités, consommations exclusives les unes des autres, etc...) et celles du consommateur.

La révolution parétienne des années 30, initiée notamment par les travaux de Roy Allen (1906) et de John Hicks (1904-1989) a définitivement assis la vision ordinaliste de l'utilité. Comme l'avaient indépendamment démontré, au tournant du siècle, Vilfredo Pareto (1848-1923) et Henri Poincaré (1854-1912), il est en effet possible de dériver des seules préférences des indices d'utilité capables de déterminer le score de chaque panier.

Dans cette approche, les préférences définies sur les consommations de X sont introduites et représentées par la relation $\succeq$ [15]. Cette relation est

[15] **Notations :** $\succeq$ est une préférence faible, $\succ$ une préférence stricte, $\sim$ une indifférence : $x \sim y$ signifie que l'agent est indifférent entre le panier x et le panier y, $x \succ y$ que l'agent préfère strictement x à y, $x \succeq y$ que l'agent préfère x à y, i.e qu'il le préfère strictement ou est indifférent entre eux. On a donc :

$$\begin{array}{rcl} x & \succeq & y \Leftrightarrow x \succ y \text{ ou } x \sim y \\ x & \sim & y \Leftrightarrow x \succeq y \text{ et } y \succeq x \\ x & \succ & y \Leftrightarrow x \succeq y \text{ et } y \not\succeq x \end{array}$$



Figure~16: Roy Allen et John Hicks

généralement supposée être:

- *réflexive*:
$$\forall x \in X,\ x \succeq x$$
- *transitive*[16]: $\forall x,\ y,\ z \in X : x \succeq y,\ y \succeq z \Rightarrow x \succeq z$

Elle constitue donc un pré-ordre qui au surplus est *complet*: [17]

$$\forall x, y \in X : x \succeq y \ ou\ y \succeq x$$

Une des étapes essentielles de la microéconomie, pour pouvoir appliquer les outils les plus courants des mathématiques est de démontrer que les préférences peuvent être représentées par des fonctions d'utilité, i.e des fonctions vérifiant la définition suivante:

ou $x \succ y \Leftrightarrow x \succeq y$ et $x \nsim y$

[16]La transitivité est une hypothèse intuitive assurant une certaine cohérence des choix. Cependant, on doit souligner deux points.

D'une part, cette propriété n'est nullement nécessaire pour développer une théorie de la demande et plus généralement une théorie de l'équilibre des marchés. Au prix d'un renforcement d'autres propriétés, notamment la convexité des préférences, les principaux résultats de la théorie de l'équilibre général demeurent en effet.

D'autre part, la transitivité n'est pas une propriété (ou une manifestation) de la rationalité des agents économiques. En effet, la transitivité est une propriété des seules préférences et est sans rapport avec le problème de l'adéquation des moyens aux fins.

[17]Autrement dit, sur tout couple de panier qu'on lui demande de comparer, l'agent a une opinion.

Définition 1 *Une fonction $U : X \to \Re$ est une fonction d'utilité si pour tout couple $(x, y) \in X$:*

$$x \succeq y \Leftrightarrow U(x) \geq U(y)$$

$$x \succ y \Leftrightarrow U(x) > U(y)$$

Une fonction d'utilité est une fonction préservant l'"ordre" des préférences sur les paniers de biens. Elle constitue donc une mesure *ordinale* des préférences, et seulement cela. N'étant qu'un indice qui range les paniers, elle n'est donc pas l' "*hedonimeter*" dont rêvait Edgeworth: la seule information économique que révèle par exemple $U(x) = 2U(y)$ (> 0) est une préférence stricte pour le panier x. En aucun cas, une telle relation n'implique que "la statisfaction procurée par le panier x est deux fois plus intense que celle procurée par le panier y": l'indice d'utilité est une mesure ordinale et non cardinale.

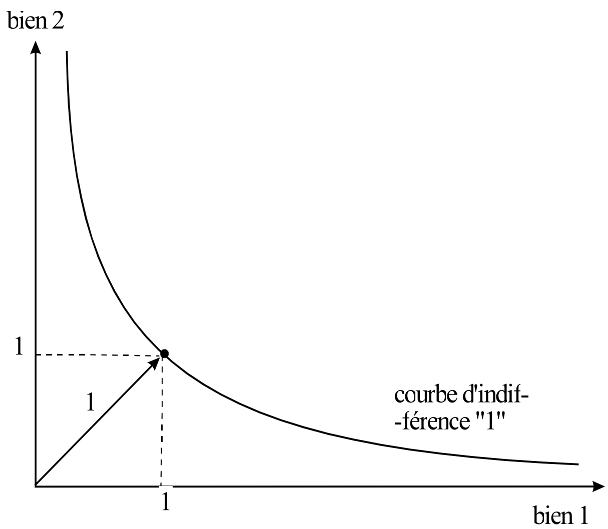
Un des premiers résultats d'existence fut obtenu par le statisticien-économiste scandinave Herman Wold (1908-1992) en 1943 dans une série d'articles importants [Wol43]. Par la suite, des résultats supplémentaires sur l'existence de telles fonctions d'utilité ont été obtenus, et notamment par John von Neumann (1903-1957) en 1944 [VNM44], par Gérard Debreu (1921-) [Deb64].[18] Sous une hypothèse de continuité des préférences, Debreu a notamment montré que tout pré-ordre complet pouvait être représenté par une fonction d'utilité *continue*. Sans qu'il soit question d'aborder cette démonstration, il est sans doute utile d'en restituer l'intuition. Pour l'illustrer, nous nous référerons à la construction (plus ancienne) utilisée par Wold pour démontrer le théorème d'existence suivant:

Théorème 1 (Wold (1943)) *Si les préférences sont définies sur $\Re^L_+$, complètes, réflexives, transitives, continues et fortement monotoniques alors il existe une fonction d'utilité représentant ces préférences.*

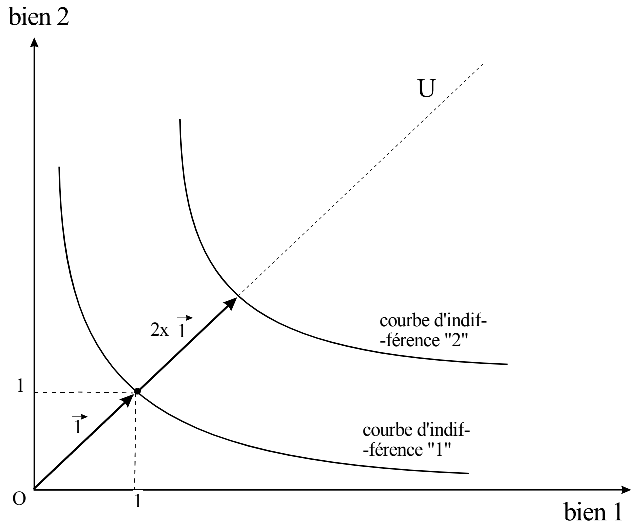
L'intuition de la démonstration et la construction de la fonction d'utilité sont illustrées sur les figures 18, 19, 20. La méthode consiste à sélectionner un panier, par exemple le panier **1**, le vecteur dont toutes les composantes sont égales à 1 (figure 18). La courbe d'indifférence de ce panier est renommée courbe "1". Grâce au panier étalon **1**, on peut également renommer d'autres courbes d'indifférences: ainsi, comme sur la figure 19, la courbe d'indifférence définie par le panier double du panier **1** est la courbe d'indifférence

[18] Une synthèse de l'ensemble de ces résultats est présentée par Peter Fishburn dans son ouvrage [Fis70].

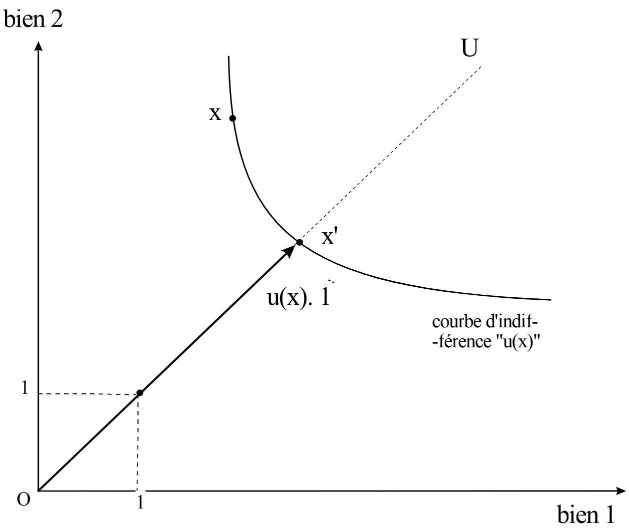
Figure~17: Wold



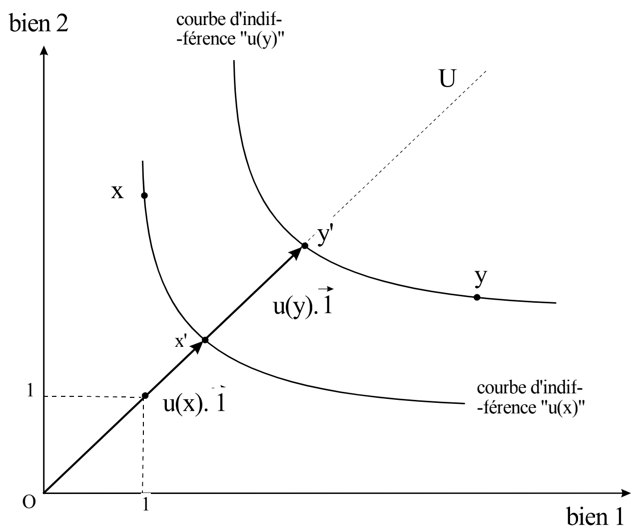
Figure~18: Le panier étalon de la fonction d'utilité



Figure~19: Utilisation du panier étalon pour renommer les courbes d'indifférence



Figure~20: Détermination de l'utilité d'un panier arbitraire x.



Figure~21: La fonction u() respecte les préférences

2. Ce nouvel indiçage définit en fait l'utilité des paniers. Cette procédure se généralise à l'ensemble des paniers. Ainsi, sélectionnons un panier arbitraire x représenté sur la figure 20. x définit une unique courbe d'indifférence qui, en raison de l'hypothèse de monotonicité des préférences, coupe nécessairement la droite OU. Appelons x' le panier défini par l'intersection de la courbe d'indifférence définie par le panier x et la droite OU. Les paniers x et x' sont équivalents pour l'agent considéré: $x' \sim x$. Le vecteur $\overrightarrow{Ox'}$ est évidemment colinéaire au panier étalon $\mathbf{1}$ et donc il existe un unique *nombre* $U(x)$ vérifiant: $\overrightarrow{Oy} = U(x).\mathbf{1}$. Ce nombre vérifie la relation d'indifférence suivante:

$$x \sim U(x).\mathbf{1}$$

puisque $x' \sim x$. Comme pour tout panier x appartenant à l'ensemble de consommation X il existe un tel $U(x)$, on vient de construire par cette procédure une fonction:

$$U : X \rightarrow \Re$$

Les propriétés de continuité des préférences assurent en outre que cette fonction est continue.

La fonction U a comme propriété essentielle de respecter les préférences. En effet, pour tout couple de panier x et y appartenant à l'ensemble de consommation et vérifiant par exemple:

$$y \succeq x$$

alors nécessairement, par construction, les réels $U(x)$ et $U(y)$ vérifient:

$$U(x).\mathbf{1} \sim x,\ U(y).\mathbf{1} \sim y$$

La transitivité des préférences assure que:

$$U(y).\mathbf{1} \sim y \succeq x \sim U(x).\mathbf{1} \Rightarrow U(y).\mathbf{1} \succeq U(x).\mathbf{1}$$

Comme les vecteurs $U(y).\mathbf{1}$ et $U(x).\mathbf{1}$ sont colinéaires, nécessairement on doit avoir:

$$U(y) \geq U(x)$$

En effet, si l'on avait $U(y) < U(x)$, la monotonocité des préférences imposerait que:

$$U(y).\mathbf{1} \prec U(x).\mathbf{1}$$

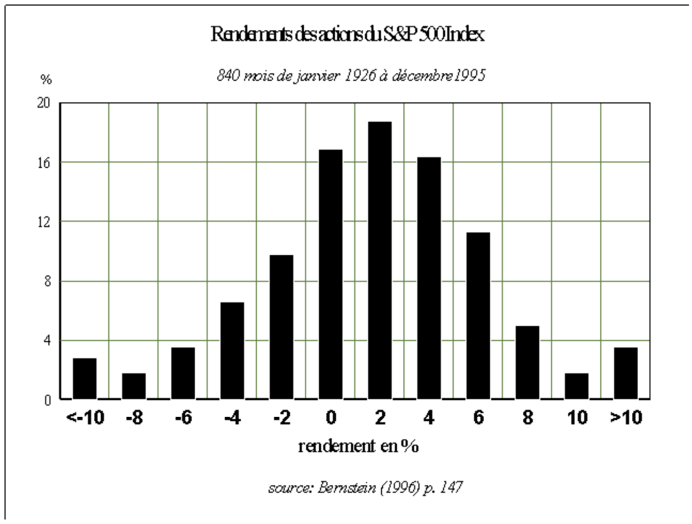
ce qui contredirait $U(y).\mathbf{1} \succeq U(x).\mathbf{1}$. La fonction U respecte bien le préordre des préférences. Elle est donc bien une fonction d'utilité.

Cette construction "pragmatique" (et dénuée de toute métaphysique) est également celle utilisée pour obtenir des fonctions d'utilité dans l'incertain.

3.2 Loteries et choix dans l'incertain

Les décisions dans l'incertain sont des des décisions d'investissement, d'épargne, de portefeuille, etc... Les revenus des futurs investissements, les dividendes que rapportent les actifs financiers sont souvent aléatoires. La figure 58 représente ainsi la distribution *effective* du rendement mensuel des actions du S&P 500 entre 1926-1995. Comme le suggère ce graphique, ces rendements peuvent (sembler) respecter certaines structures, voire être des distributions normales. Aussi, si l'ensemble des objets que l'on doit choisir est fort divers, formellement sa structure n'est pas sans évoquer celle des loteries: des gains possibles caractérisés par des probabilités d'occurence. Au surplus, avec les loteries comme avec les actifs financiers: "l'argent que les joueurs ont mis dans le jeu ne leur appartient plus [...] mais [...] ils reçoivent en contrepartie le droit de recevoir ce que la chance leur accordera, selon les règles qu'ils ont accepté au début du jeu." (Blaise Pascal, cité par [Ber96] p. 67)

Aussi, la théorie de la choix dans l'incertain, développée par von Neumann et ses continuateurs immédiats (Marschak [Mar50], Samuelson [Sam52], Herstein & Milnor [HM53]), suppose que les objets sur lesquels sont définies les préférences sont des loteries. Ceci est cependant restrictif. Une loterie est caractérisée par une distribution de probabilité *objective* des prix. Assimiler les objets à des loteries revient donc à supposer qu'aux événements possibles correspondent des probabilités "objectives" sur lesquelles s'accordent les agents.



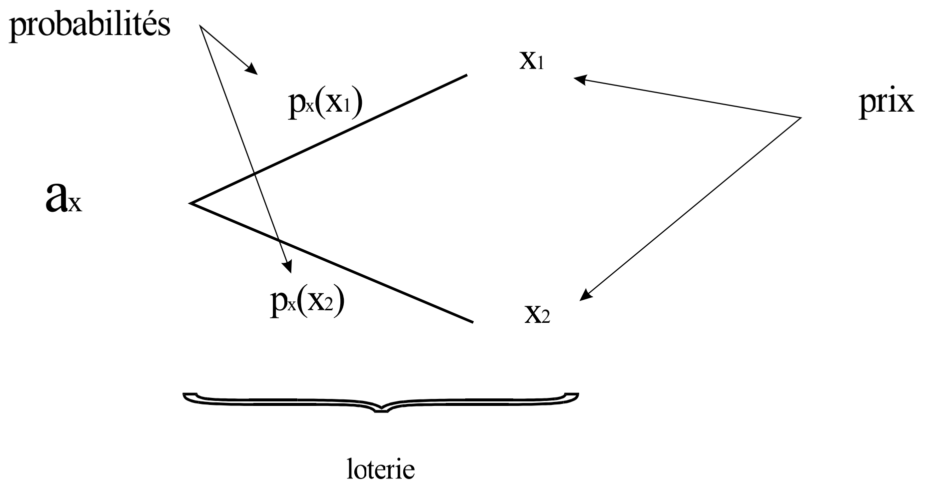
Figure~22: La distribution du rendement mensuel des actions aux Etats-Unis, 1926-1995

Les probabilités précèdent donc ici les préférences.[19] Avant d'aller plus avant, introduisons quelques notations et définitions.

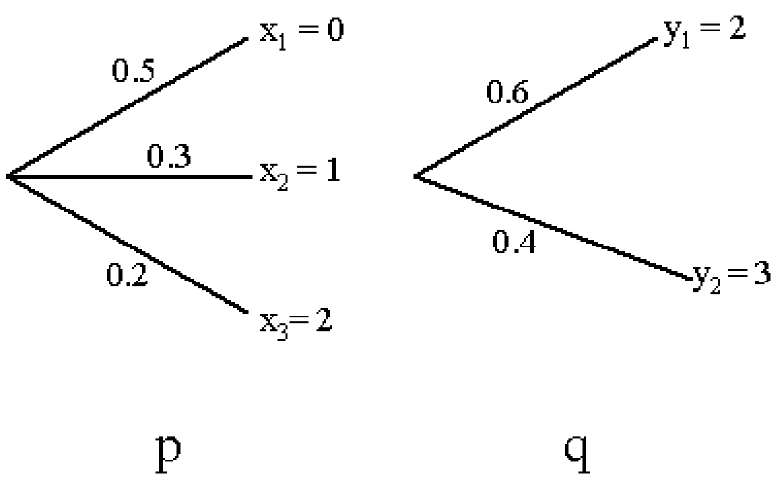
Comme l'illustre la figure 23, chaque loterie $a_x = [\overrightarrow{x}, \overrightarrow{p}_x]$ est composée de deux séries d'éléments: les différents 'prix' possibles $(x_1,\ x_2)$ et leurs probabilités $(p_x(x_1),\ p_x(x_2))$. Lorsque la loterie est une *loterie simple*, les prix sont des biens, des paniers de biens, du bien numéraire. Pour utiliser les notations déjà vues dans le certain, on notera alors X l'ensemble des prix possibles, avec $X \subset \Re^N$. L'ensemble des probabilités définies sur cet ensemble X sera noté $\Delta(X)$. L'ensemble des loteries simples est donc $X \times \Delta(X)$.

A côté de ces loteries "simples", on peut considérer des *loteries composées.* Une loterie composée est une loterie dont certains prix sont eux-mêmes des loteries, i.e. certains gains possibles de la loterie composée est de gagner le droit (et l'obligation) de participer à une autre loterie. Ainsi sur la figure **??**, la loterie composée permet de recevoir avec une probabilité $1-q$ un panier y. Par contre, avec une probabilité p, le prix reçu est la loterie $[(x, y), (p, 1-p)]$. Ex ante, participer à la loterie composée donne droit à recevoir au total le

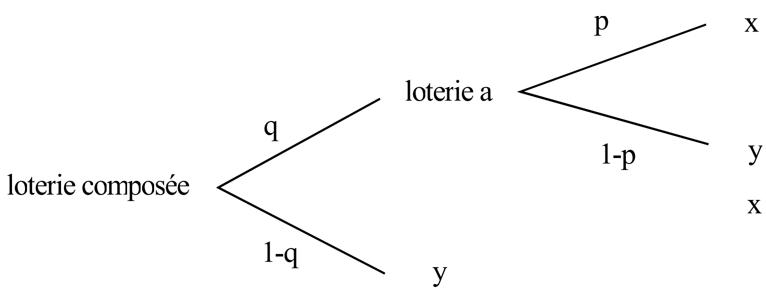
[19]Une approche alternative fut proposée très tôt par Léonard Savage. Dans celle-ci, les données objectives se réduisent aux événements. A priori, il n'existe pas nécessairement de probabilités "objectives". Les agents peuvent donc ne pas avoir les mêmes croyances sur les probabilités de réalisation des différents événements. Ces probabilités sont, comme les préférences, des éléments subjectifs définissant chaque personnalité. Cependant, ce relâchement considérable du modèle de décision ne modifie pas considérablement les réprésentations obtenues. Aussi, en restons-nous essentiellement aux probabilités objectives.



Figure~23: Représentation d'une loterie



Figure~24:



Figure~25: Un exemple de loterie composée

prix x avec une probabilité totale qp, le prix y avec une probabilité $q(1-p)+(1-q)=1-qp$. Comme $q \in [0,1]$, le droit de recevoir y peut être assimilé à une loterie $b=[(y,0),(1,0)]$ donnant avec certitude y, la loterie composée n'est qu'une combinaison convexe de loteries que l'on notera $q * a \oplus (1-q) * b$. Comme la distribution des prix x et y est similaire à celle de la loterie $[(x,y),(qp,1-qp)]$, une telle combinaison convexe est elle-même une loterie. Cette propriété des loteries composées assure que l'ensemble des loteries, que l'on notera $\pounds$, est un ensemble convexe:

$$\forall a_x, a_y \in \pounds,\ \forall \alpha \in [0,1] : \alpha * a_x \oplus (1-\alpha) * a_y \tag{3}$$

Remarque 1 *Comme l'illustre la figure 26, il est toujours possible de représenter une loterie (comme la loterie centrale du graphique) par plusieurs loteries composées.*

Les préférences (sur les loteries) étant toujours représentées par les sigles $\succeq$, $\succ$, $\sim$, il est nécessaire d'introduire quelques hypothèses préliminaires, dont certaines sont purement formelles comme les deux suivantes:

Hypothèse 1 (L 1) $\forall a_x, a_y \in \pounds,\ 1 * a_x \oplus 0 * a_y \sim a_x$

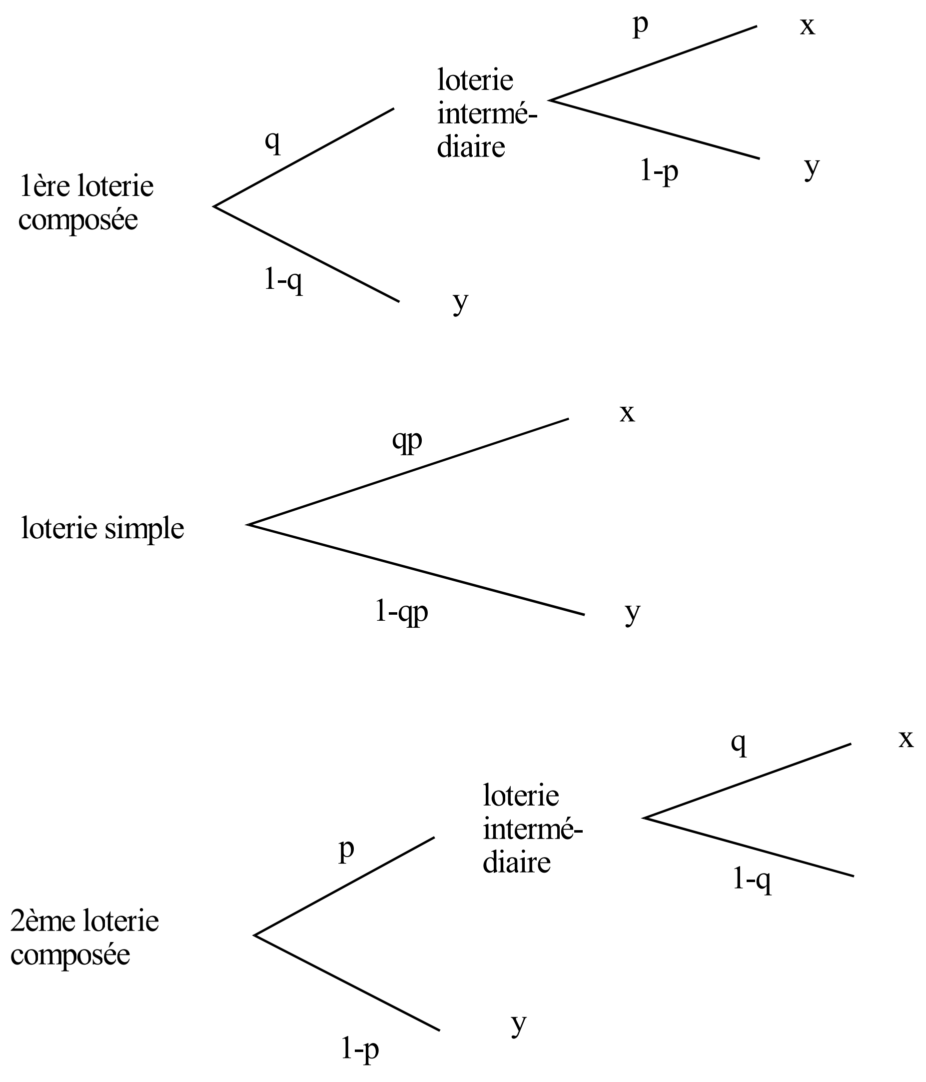
Hypothèse 2 (L 2) $\forall a_x, a_y \in \pounds$, $\forall \alpha \in [0,1]$:

$$\alpha * a_x \oplus (1-\alpha) * a_y \sim (1-\alpha) * a_y \oplus \alpha * a_x$$

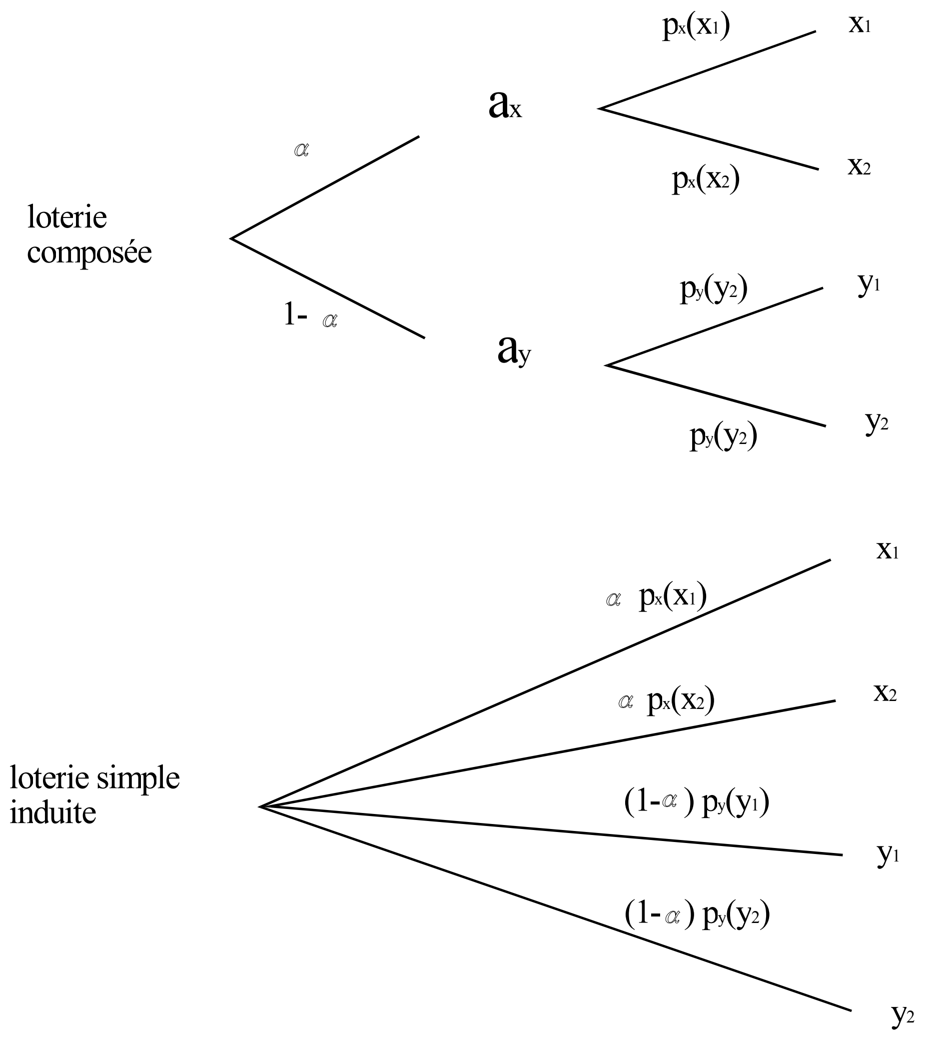
Une troisième, appelée *réduction des loteries composées*, revient à supposer que les loteries composées et les loteries simples qu'elles définissent sont équivalentes pour les agents (cf figure 27):

Hypothèse 3 (L 3) $\forall a_x = [\overrightarrow{x}, \overrightarrow{p}_x] \in \pounds$, $\forall a_y = [\overrightarrow{y}, \overrightarrow{p}_y] \in \pounds$, $\forall \alpha \in [0,1]$:

$$\alpha * a_x \oplus (1-\alpha) * a_y \sim [(\overrightarrow{x}, \overrightarrow{y}), (\alpha p_x, (1-\alpha) p_y)]$$



Figure~26: Représentation d'une loterie simple au centre sous forme de plusieurs loteries composées.



Figure~27: Une loterie composée et sa loterie simple équivalente

Si l'on suppose que les préférences constituent un pré-ordre complet $\succeq$ sur les loteries, que les préférences sont continues, il est possible d'appliquer immédiatement les théorèmes généraux d'existence des fonctions d'utilité (comme celui de Debreu dans [Deb59]) pour démontrer l'existence d'une fonction continue U:

$$\begin{aligned} U &: \pounds \to \Re \\ a_x &= (x, p_x) \mapsto U = U(a_x) = U(x; p_x) \end{aligned}$$

représentant le pré-ordre $\succeq$:

$$\forall a_x, a_y \in \pounds,\ a_x \succ a_y \Leftrightarrow U(a_x) > U(a_y)$$

Ce type de représentation très générale suffit pour l'analyse de nombreux problèmes. En particulier, l'extension de la théorie de l'équilibre général à l'incertain ne nécessite aucune restriction supplémentaire. Cependant, dans les domaines de l'assurance, de la finance, certaines restrictions s'avèrent utiles et fructueuses. Comme l'avait déjà noté Bernoulli, la représentation des préférences sous forme d'*utilité espérée* permet d'obtenir de nombreux résultats intéressants. Les fonctions d'utilité U s'écrivent alors sous la forme additive suivante:

$$\forall a_x = \left[(x_i)_{i=1,\ldots}, (p_x(x_i))_{i=1,\ldots} \right],\ U(a_x) = \sum_i p_x(x_i) .u(x_i) \qquad (4)$$

où u est la fonction d'utilité élémentaire (appelée aussi fonction vNM en l'honneur de von Neumann et Morgenstern). Pour obtenir ceci, des axiomes supplémentaires sont introduit dont l'*axiome d'indépendance* suivant:

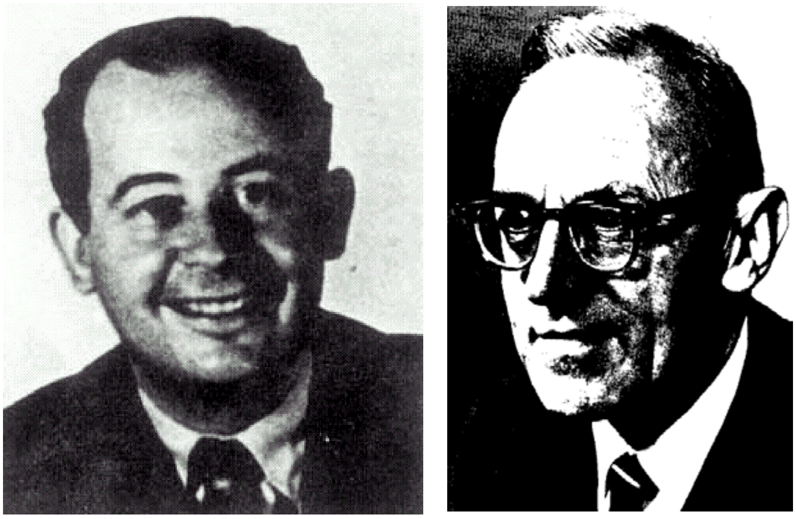
Axiome 1 *Soient deux loteries x, $y \in \pounds$ vérifiant:*

$$x \succ y$$

alors pour toute loterie $z \in \pounds$, pour tout réel $\alpha \in]0, 1[$ on a:

$$\alpha * x \oplus (1 - \alpha) * z \succ \alpha * y \oplus (1 - \alpha) * z$$

Mais, la construction axiomatique menant à l'utilité espérée étant relativement aride, il est sans doute sage de sursoir provisoirement à son exposé pour analyser maintenant ses nombreux apports.



Figure~28: John von Neumann et Oskar Morgenstern

4 Aversion à l'égard du risque: mesures et conséquences

Dans l'incertain, les agents économiques sont souvent amenés à prendre des décisions d'investissement, de portefeuille plus ou moins risqués. Ces décisions sont évidemment déterminées pour partie par les données du marché (prix, taux d'intérêt), pour partie par certaines variables objetives caractérisant ces agents (richesse, âge notamment). Mais, intuitivement, on comprend bien que ces choix dépendent aussi de manière fondamentale de paramètres subjectifs, des préférences et des goûts des décideurs, de leur attitude à l'égard du risque: toutes choses égales par ailleurs, notamment quant à leurs richesses, il semble évident que le risque de leurs positions financières sera d'autant plus faible qu'ils seront "prudents" (=?), qu'ils auront de l'aversion à l'égard du risque (=?). Mais *peut-on définir rigoureusement, dans le cadre d'un modèle théorique, ces notions?*

L'objet de cette section est d'exposer les notions qui ont été développées notamment dans le cadre de l'utilité espérée pour représenter ces attitudes et de montrer leurs applications à la demande d'actifs financiers et à la demande d'assurance.

4.1 Variables aléatoires, états du monde

Les économies représentées tout au long de cette section sont caractérisées par une ou plusieurs sources d'incertitude, i.e. il existe un ensemble *fini* d'événements *dont les probabilités sont données*. Formellement, ces événements sont des variables aléatoires $\widetilde{f}$ résumées par leurs valeurs possibles $\{f_1, f_2, ..., f_m\}$ et par les probabilités de ces différentes valeurs: $\left\{\pi_1^f, ..., \pi_m^f\right\}$. Par exemple, comme sur la partie supérieure de la figure 30, on peut avoir des aléas portant sur la météo; un ensemble de 4 événements:

$\{soleil, pluie, avec\ vent, sans\ vent\}$ est ainsi spécifié; les probabilités de "soleil" et "pluie" sont respectivement $1/3$ et $2/3$; la probabilité ($totale$) de l'événement "avec vent" est $\frac{1}{3}\frac{1}{2} + \frac{2}{3}\frac{1}{2} = 1/2$ comme celle de l'événement "sans vent": $\frac{1}{3}\frac{1}{2} + \frac{2}{3}\frac{1}{2} = 1/2$. Ces événements ont comme propriété de ne pas être nécessairement exclusifs les uns des autres: comme le montre l'arbre des événements, on peut par exemple avoir simultanément les événements "soleil" et "avec vent", ou encore les événements "pluie" et "sans vent". Comme dans cet exemple, il n'existe cependant que deux sources d'incertitude (illustrée par les enchaînements graphiques des événements), formellent, ce cadre peut donc être représenté par la donnée de deux variables aléatoires:

- la variable aléatoire $\widetilde{g}$ dont les valeurs sont $\{g_1, g_2\}$ avec:

$$g_1 := \ soleil$$

$$g_2 := \ pluie$$

  dont les probabilités sont:

$$\pi_1^g = \frac{1}{3},\ \pi_2^g = \frac{2}{3}$$

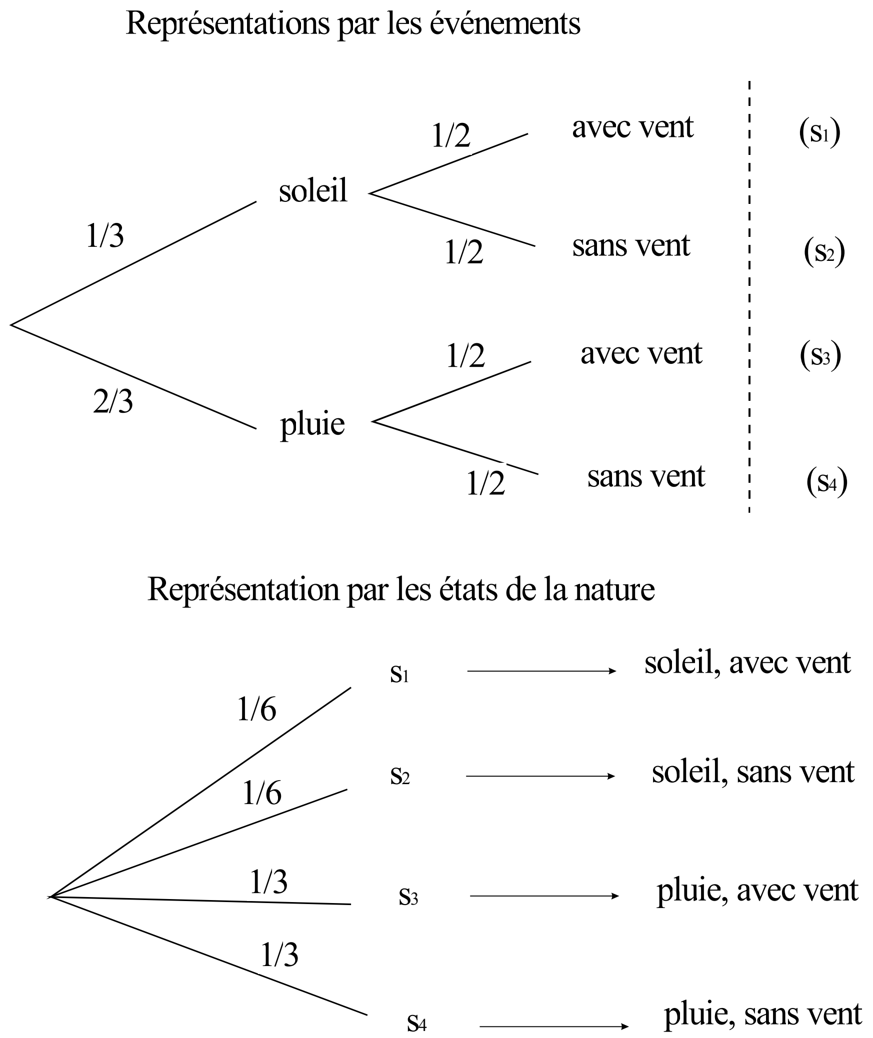
- la variable aléatoire $\widetilde{h}$ dont les valeurs sont $\{h_1, h_2\}$ avec:

$$h_1 := \ avec\ vent$$

$$h_2 := \ sans\ vent$$

  dont les probabilités sont:

$$\pi_1^h = \frac{1}{2},\ \pi_2^h = \frac{1}{2}$$



Figure~29: Un exemple de représentation de l'incertain par un ensemble d'états de la nature

Figure~30: Andrei Kolmogorov, l'Euclide de la théorie des probabilités: "La théorie des probabilités est une discipline mathématique qui peut et doit se développer à partir d'axiomes exactement de la même façon que la géométrie ou l'algèbre" (extraits de ses Fondements de la Théorie des Probabilités, 1933)
A gauche, à l'époque des Fondements. A droite en 1964 lors de son entrée à la Royal Society de Londres.

Lorsque l'on multiplie cependant le nombre de sources, i.e. le nombre d'embranchements, cette représentation en terme d'événements peut devenir formellement fort lourde. Aussi, depuis la révolution initiée en théorie des probabilités par Andrei Kolmogorov (1903-1987), l'Euclide de la théorie des probabilités, on lui préfère une représentation plus abstraite mais plus compacte en terme d'états de la nature (ou états du monde). L'incertitude est alors représentée par un ensemble S d'états de la nature:

$$S = \{s_1, s_2, ..., s_n\}$$

dont les probabilités sont notées:

$$\pi = (\pi(s_1), \pi(s_2), ..., \pi(s_n))$$

ou encore:

$$\pi = (\pi_1, \pi_2, ..., \pi_n)$$

L'état sélectionné par Dame Nature détermine alors la valeur des autres variables f. Comme l'illustre la figure 30, une méthode très simple pour construire cet ensemble d'états de la nature consiste simplement à renommer les extrémités de l'arbre (s_1, s_2, s_3, s_4 à droite du 1er arbre), à calculer la probabilité de ces extrémités, et à substituer à cet arbre un second définies uniquement sur les extrémités du précédent $\{s_1, s_2, s_3, s_4\}$ dont les probabilités sont donc:

$$\pi(s_1) = \frac{1}{6},\ \pi(s_2) = \frac{1}{6},\ \pi(s_3) = \frac{1}{3},\ \pi(s_4) = \frac{1}{3}$$

Les états de la nature déterminent alors les autres variables. Ainsi, les variables aléatoires deviennent des fonctions définies sur les états:

$$\begin{aligned} f &: S \rightarrow \{soleil,\ pluie\} \\ f(s_1) &= f(s_2) = soleil \\ f(s_3) &= f(s_4) = pluie \end{aligned}$$

$$\begin{aligned} g &: S \rightarrow \{avec\ vent,\ sans\ vent\} \\ g(s_1) &= g(s_3) = avec\ vent \\ g(s_2) &= g(s_4) = sans\ vent \end{aligned}$$

La distribution de chaque variable aléatoire est alors définie par celle des états du monde; ainsi par exemple:

$$\pi^f(soleil) = \sum_{s \in \{s : f(s) = soleil\}} \pi(s)$$

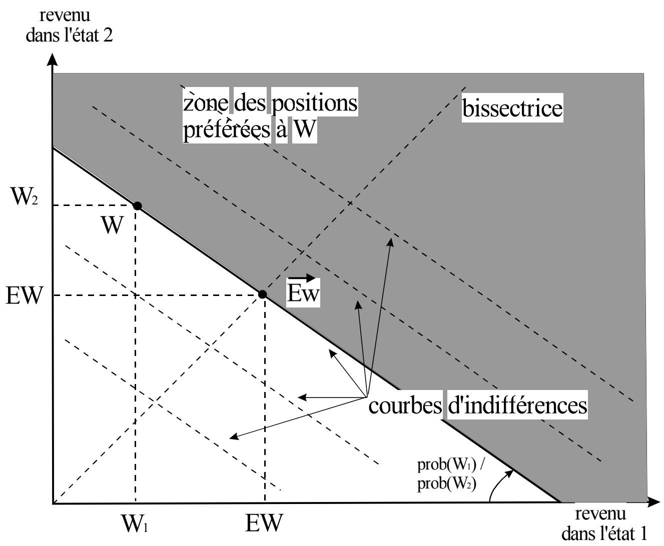
Cette méthode de construction permet donc d'engendrer les états de la nature connaissant les événéments et leurs enchaînements lorsque les événements sont en nombre fini (ou dénombrables). Mais cette construction se généralise à des continuums également.

Pour simplifier, on supposera que dans chaque état s, les variables sur lesquelles sont définies partiellement les préférences peuvent être résumées par un scalaire, le revenu par exemple, que l'on notera w_s (ou encore $w(s)$). Les préférences sont représentées par une fonction d'utilité U:

$$\begin{aligned} U &: \Re_+^n \times \Delta(S) \rightarrow \Re \\ (w_1, ...w_n; \pi_1, ..., \pi_n) &\mapsto U(w_1, ...w_n; \pi_1, ..., \pi_n) \end{aligned}$$

Parfois, on renforcera une représentation sous forme d'utilité espérée:

$$U(w_1, ...w_n; \pi_1, ..., \pi_n) = \sum_{i=1}^{n} \pi_i . u(w_i)$$



Figure~31: Les positions préférées par un agent neutre au risque possédant initialement W

Dans les deux cas, on supposera aussi que l'utilité est strictement croissante des revenus:

$$\forall w \in \Re_+^n,\ \frac{\partial}{\partial w_i} U\left(w_1, ...w_n; \pi_1, ..., \pi_n\right) > 0,\ i = 1, ..., n$$

$$\forall w \in \Re_+,\ \frac{\partial}{\partial w} u\left(w\right) > 0$$

4.2 L'aversion à l'égard du risque

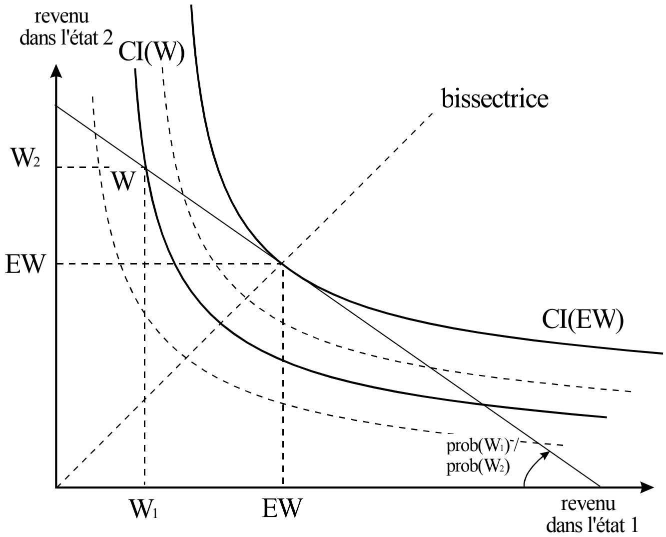
Supposons que $\#S = 2$, que l'individu considéré ait des préférences équivalentes à la fonction:

$$U = \pi_1.w_1 + \pi_2.w_2$$

et possède intialement le profil de revenu $w = (w_1, w_2)$ représenté sur la figure **??**. Un lieu particulièrement intéressant est la bissectrice: celle-ci est l'ensemble des revenus $z = (z_1, z_2)$ où $z_2 = z_1 = \mathbf{E}z$ - $\mathbf{E}z$ l'espérance de z, i.e. l'ensemble des profils de revenu sans risque. Une propriété remarquable des préférences est évidemment que l'on a nécessairement (comme $\pi_1 + \pi_2 = 1$):

$$\pi_1\left(\mathbf{E}w_1\right) + \pi_2\left(\mathbf{E}w_2\right) = \mathbf{E}w = \pi_1.w_1 + \pi_2.w_2 = U$$

Un tel individu est donc indifférent à détenir w ou à détenir la moyenne de w, i.e. $(\mathbf{E}w, \mathbf{E}w)$. Plus largement, il sera également indifférent entre (z_1, z_2)



Figure~32: La préférence pour la moyenne lorsque les préférences sont convexes.

et w dès lors que:

$$\pi_1.z_1 + \pi_2.z_2 = \mathbf{E}w$$

Partant de w, les variations de revenu (dz_1, dz_2) gardant constante l'utilité sont:

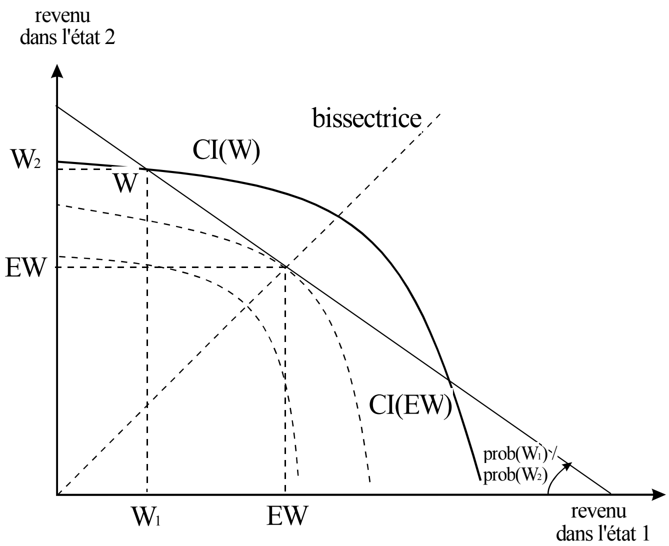
$$\pi_1.dz_1 + \pi_2.dz_2 = 0 \Rightarrow -\frac{dz_2}{dz_1} = \frac{\pi_1}{\pi_2}$$

Les courbes d'indifférence d'un tel individu sont des droites parallèles les unes aux autres de pente π_1/π_2. L'ensemble des positions acceptables par l'individu, i.e. lui rapportant une utilité supérieure à w, est le demi-plan supérieur à la droite $(w, \mathbf{E}\overrightarrow{w})$, où $\mathbf{E}\overrightarrow{w}$ est le vecteur dont toutes les composantes sont égales à l'espérance $\mathbf{E}w$.

Lorsque l'individu a des préférences quelconques U, il n'est plus nécessairement indifférent entre w et $\overrightarrow{\mathbf{E}w}$. Ainsi, avec les courbes d'indifférences de la figure [**?**], la position sûre $\overrightarrow{\mathbf{E}w}$ est strictement préférée à la position initiale. Si les courbes d'indifférence sont strictement convexes, il est aisé de voir qu'en tout point w n'appartenant pas à la bissectrice, une telle préférence pour la sûreté émergera également:

$$\forall w \neq \overrightarrow{\mathbf{E}w},\ \overrightarrow{\mathbf{E}w} \succ w \ \textit{ou encore}\ U\left(\mathbf{E}w;\pi_1,\pi_2\right) > U\left(\mathbf{E}w;\pi_1,\pi_2\right)$$

Cette préférence universelle pour la sûreté révèle donc une *aversion à l'égard du risque*, une *risquophobie* de l'individu. Le revenu certain EC pour



Figure~33: La préférence pour le risque lorsque les préférences sont strictement concaves.

lequel il est indifférent à w:

$$EC \;/\; EC \sim w$$

que l'on appelle *équivalent certain*, est alors inférieur à l'espérance comme l'illustre la figure 34.

A contrario, lorsque les préférences sont strictement concaves (figure 33), l'individu préférera une position risquée comme w à la position certaine définie par sa moyenne: le comportement de l'individu révèle un tempérament joueur, une risquophilie.

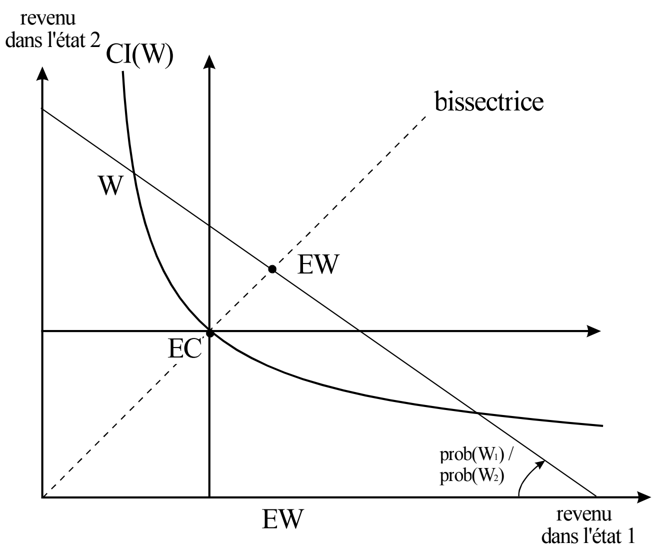
La considération de la structure convexe (ou non convexe) des préférences permet donc de représenter la notion d'aversion à l'égard du risque. Mais comparant deux individus ayant le même revenu w, peut-on ordonner leurs aversions?

Pour répondre à cette question, considérons un profil de revenu sûr w comme celui illustré sur la figure 35. Les points A, B, C, D, E sont des positions possibles; A, B sont dominés par la position w:

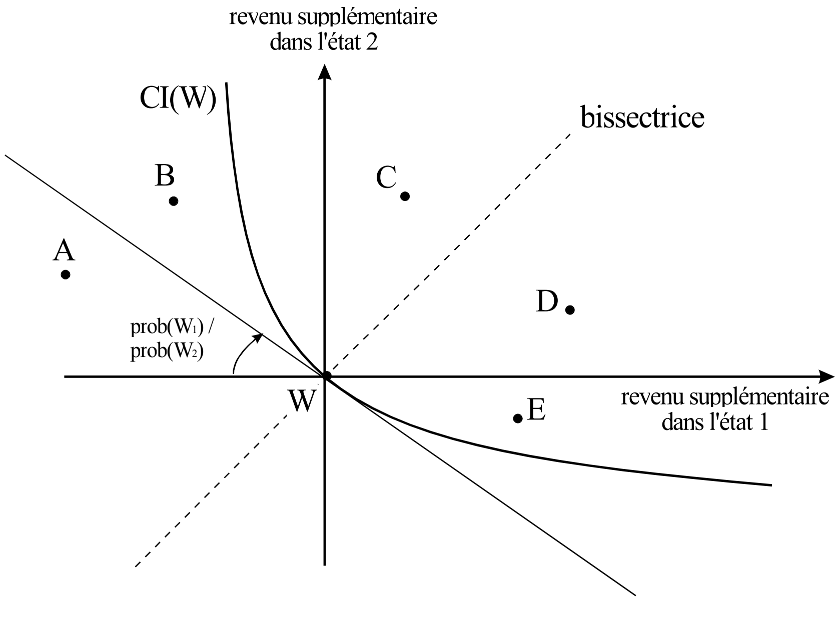
$$w \succ A, \; w \succ B$$

et les positions C, D et E sont acceptables:

$$C \succ w, \; D \succ w, \; E \succ w$$



Figure~34: Détermination de l'équivalent certain pour des préférences strictement convexes.



Figure~35: Ensemble des positions acceptables

La droite passant par $\overrightarrow{w} = (w, w)$ et de pente π_1/π_2 est l'ensemble des positions dont l'espérance est w. Elle représente aussi la courbe d'indifférence d'un agent neutre au risque. Notons $A_{\succeq}(w)$ ou $A_U(w)$ l'ensemble des positions acceptables pour un agent possédant initialement w et donc les préférences sont résumées par le pré-ordre $\succeq$ ou par l'indice d'utilité U:

$$A_{\succeq}(w) = \left\{z \in \Re_+^2 : z \succeq w\right\}$$

$$A_U(w) = \left\{z \in \Re_+^2 : U(z;\pi) \geq U(w;\pi)\right\}$$

Lorsque l'on part initialement d'une position certaine, la stricte convexité des préférences implique que l'ensemble $A_{\succeq}(w)$ des positions acceptables est inclus dans le demi-plan d'un individu neutre au risque comme l'illustre la figure 36. Ceci suggère d'ordonner, de mesurer les aversions au risque par les relations d'inclusion des ensembles $A_{\succeq}(w)$.

Considérons pour cela la figure 37. Deux agents repérés par leurs fonctions d'utilité respectives U et V, disposant de la même richesse initiale, y sont comparés. La comparaison des positions acceptables pour les deux agents est la base de la mesure ordinale de l'aversion pour le risque proposée par Menachem Yaari [Yaa69]:

Définition 2 *Un agent dont la fonction est U a une plus grande aversion à l'égard du risque qu'un agent dont la fonction est V si pour tout niveau de revenu w, l'ensemble des positions acceptables pour le premier est un sous-ensemble de l'ensemble des positions acceptables pour le second, i.e.:*

$$AR_U > AR_V \Leftrightarrow A_U(w) = A_V(w) \;\; \forall w = \mathbf{E}w \in \Re_+^2$$

Comme le montre la figure 38, dès lors que les fonctions d'utilité (ou les pré-ordres) sont ordonnées, les équivalents certains le sont aussi:

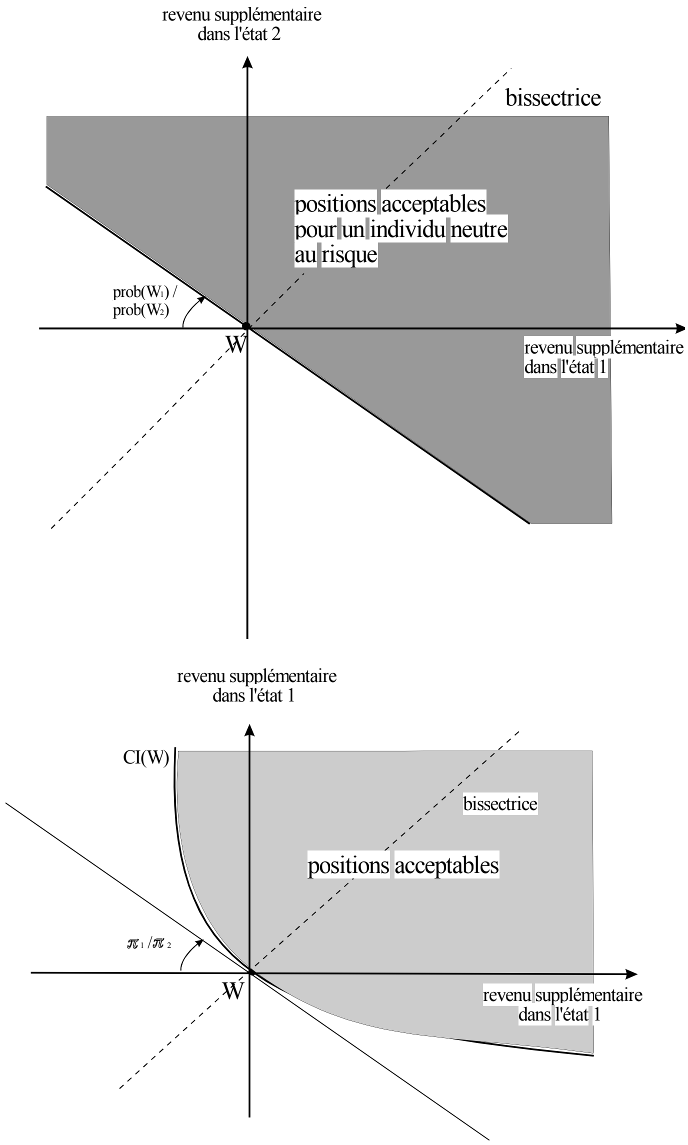
$$AR_U(w) > AR_V \Rightarrow CI_U > CI_V$$

L'individu qui est le plus risquophobe est donc prêt à consentir le sacrifice le plus important pour obtenir une position certaine. La mesure de ce sacrifice (ρ_U ou ρ_V), la *prime de risque*, est donnée par la différence entre l'espérance de la position risquée initiale w et l'équivalent certain:

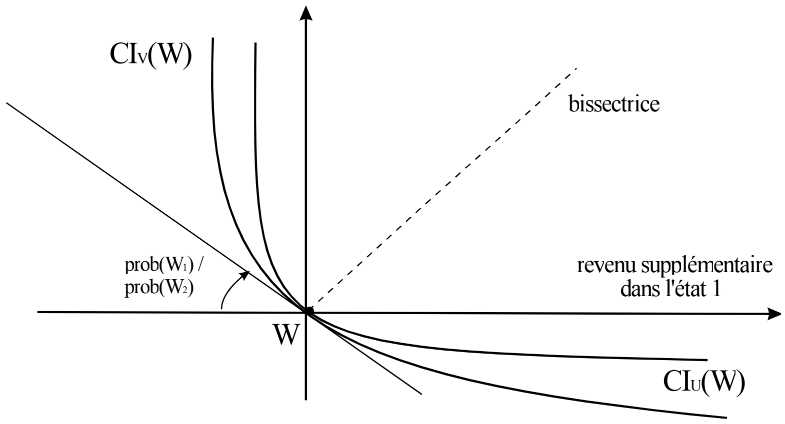
$$\rho_U = \mathbf{E}w - CI_U, \;\; \rho_V = \mathbf{E}w - CI_V$$

Par conséquent, si un individu U a une aversion au risque plus élevée qu'un agent V, sa prime de risque sera plus importante:

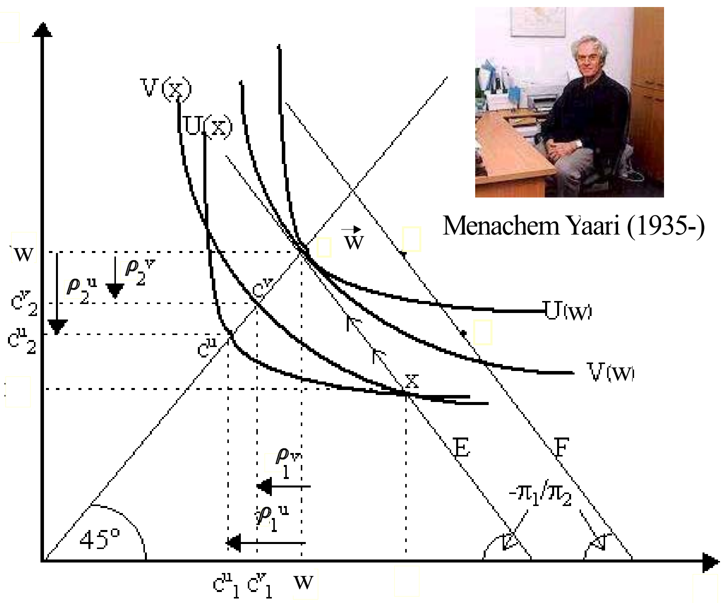
$$AR_U > AR_V \Rightarrow \rho_U(w) > \rho_V(w) \;\; \forall w$$



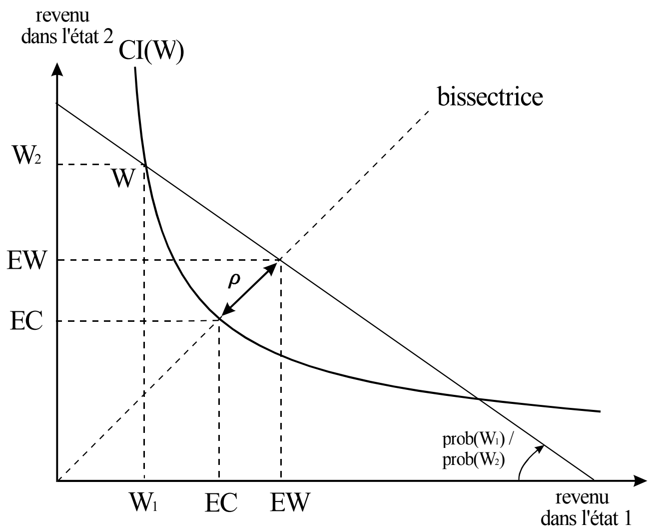
Figure~36: Comparaison des positions acceptables pour un individu neutre au risque (à gauche) et pour un individu risquophobe (à droite)



Figure~37: Comparaison deux individus U et V ayant initialement la même richesse certaine w



Figure~38: Les mesures de l'aversion au risque de Yaari (1969)



Figure~39: Détermination de la prime de risque

4.3 Utilité espérée et aversion au risque

Lorsque les fonctions d'utilité vérifient la propriété d'espérance d'utilité:

$$U = \pi_1.u(w_1) + \pi_2.u(w_2)$$

la caractérisation de l'aversion au risque peut se faire plus directement à l'aide de la fonction d'utilité u selon les voies tracées par Daniel Bernoulli en 1738.

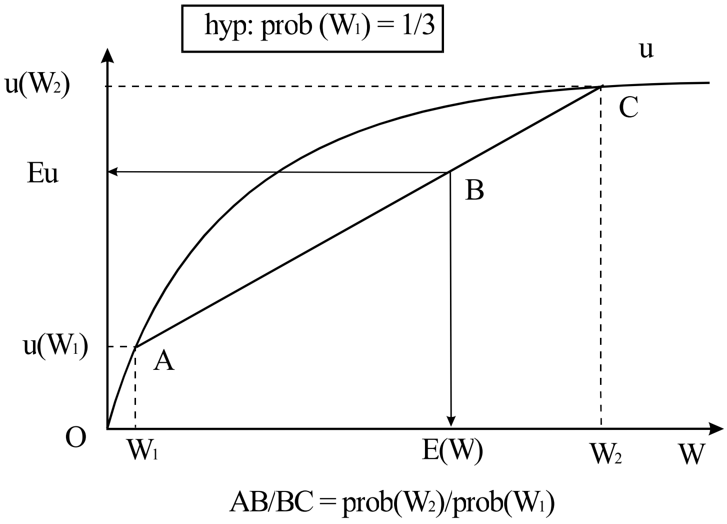
Géométriquement, lorsque la fonction d'utilité élémentaire est donnée, l'utilité espérée est déterminée par le barycentre des utilités élémentaires. Ainsi, sur la figure 40, deux niveaux de richesse sont possibles: w_1 et w_2, et donc deux niveaux d'utilité élémentaires: $u(w_1)$ et $u(w_2)$. Des points A et B ainsi définies, on détermine le barycentre B:

$$B = \pi_1.A + \pi_2.C$$

Par construction, l'abscisse de ce point est la richesse moyenne $(\pi_1.w_1 + \pi_2.w_2)$, son ordonnée est son utilité moyenne, l'utilité espérée $(\pi_1.u(w_1) + \pi_2.u(w_2))$.

Lorsque la fonction d'utilité élémentaire est strictement concave, les courbes d'indifférence de l'agent sont strictement convexes, et donc il préfère la sûreté. Ainsi, figure 41, si l'on substitue au profil (w_1, w_2) sa moyenne $\mathbf{E}w$, le gain d'utilité de l'agent est dU:

$$dU = u(\mathbf{E}w) - \pi_1.u(w_1) + \pi_2.u(w_2)$$



Figure~40: Détermination graphique de l'espérance d'utilité et du revenu espéré.

le sacrifice qu'il est prêt à consentir pour lui substituer un niveau certain dW:

$$u(\mathbf{E}w - dW) = \pi_1.u(w_1) + \pi_2.u(w_2)$$

Ce sacrifice maximum constitue la *prime de risque*.

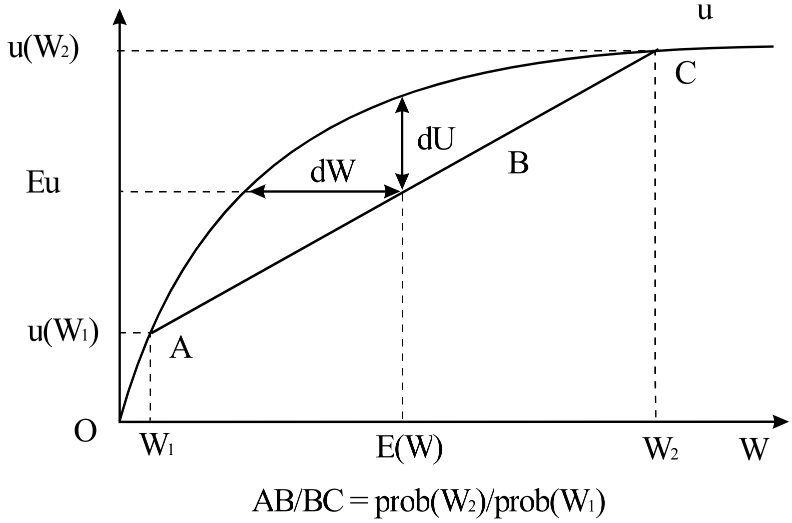
Remarque 2 *Si la fonction d'utilité était strictement convexe (figure 42), le gain d'utilité deviendrait une perte et l'agent serait prêt à payer pour obtenir une position risquée de même moyenne!*

L'hypothèse d'espérance d'utilité, par rapport à l'approche très générale de Yaari, permet cependant de donner une autre caractérisation de l'aversion à l'égard du risque. Cette caractérisation repose sur la comparaison des fonctions d'utilité élémentaires et des primes de risque.

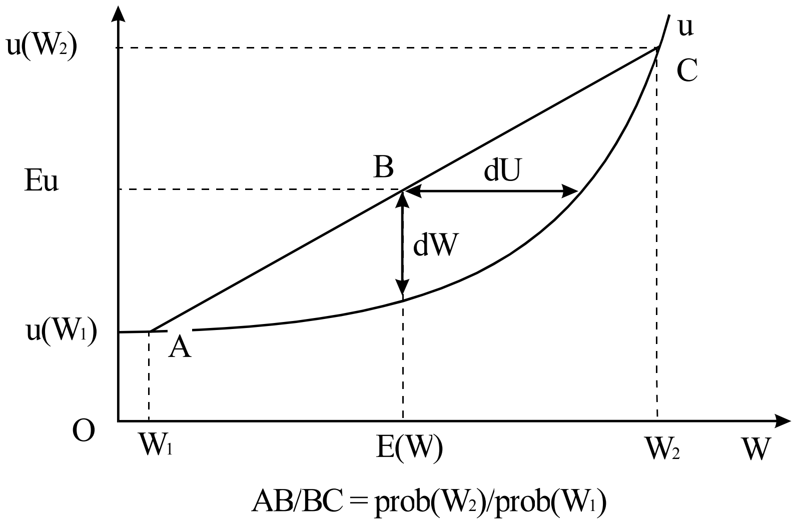
Ainsi, sur la figure 43, *partant initialement d'un niveau de richesse certain* $\overline{w}$, les richesses sont modifiées, rendues variables tout en conservant leurs niveaux moyens:

$$w_1 \neq w_2 \; mais \; \pi_1.w_1 + \pi_2.w_2 = \overline{w}$$

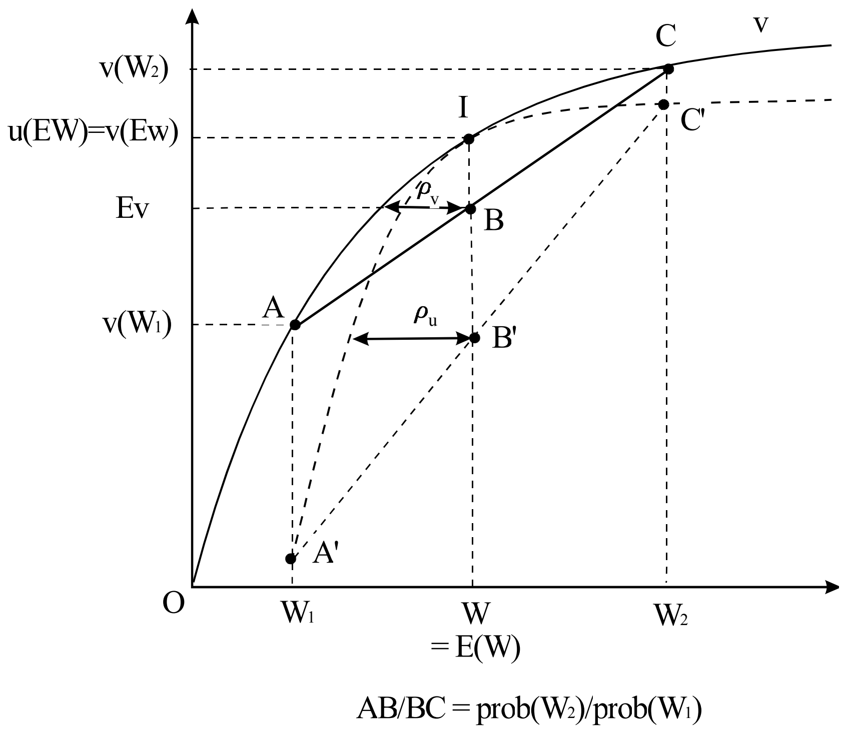
Pour une première fonction d'utilité élémentaire v (en trait plein sur la figure) sont représentées les différents niveaux d'utilité ($v(w_1)$ et $v(w_2)$), l'utilité espérée et la prime de risque ρ_v. Supposons qu'autour du point initial I on déforme la fonction v en augmentant sa courbure. On obtient alors une nouvelle fonction, notée u sur le graphique.



Figure~41: Le gain d'utilité induit par la suppression du risque et le sacrifice maximal consenti.



Figure~42: Une fonction d'utilité élémentaire d'un agent risquophile



Figure~43: Courbure de la fonction d'utilité élémentaire et prime de risque

Pour toute courbe représentant une fonction f, une mesure de sa courbure en un point est donnée par le taux de croissance instantannée de sa pente, i.e. $\left|\frac{f''}{f'}\right|$. Pour la fonction u ainsi construite on a:

$$\left|\frac{u''}{u'}\right| > \left|\frac{v''}{v'}\right|$$

ou encore:

$$-\frac{u''}{u'} > -\frac{v''}{v'}$$

puisque par leurs strictes concavités: $u'' < 0$, $v'' < 0$. Comme on peut le vérifier graphiquement, cette propriété implique que la prime de risque ρ_u de l'agent u est supérieure à la prime de risque ρ_v. Localement, ce résultat peut être retrouvé par un développement limité d'ordre 2. En effet, si l'on se situe en une richesse w (suffisamment) peu différente de la richesse moyenne $\mathbf{E}\left[\widetilde{w}\right]$, l'espérance d'utilité est approximativement: [20]

$$\mathbf{E}\left[v\left(w\right)\right] \approx v\left(\mathbf{E}\left[\widetilde{w}\right]\right) + \frac{1}{2}v''\left(\mathbf{E}\left[\widetilde{w}\right]\right).\boldsymbol{\sigma}_w^2 \tag{5}$$

où $\sigma_w^2 = \mathbf{E}\left[\left(w - \mathbf{E}\left[\widetilde{w}\right]\right)^2\right]$ est la variance de la richesse.

De la définition de la prime de risque:

$$\mathbf{E}\left[v\left(w\right)\right] = v\left(\mathbf{E}\left[\widetilde{w}\right] - \rho_v\right)$$

on en déduit le développement limité:

$$v\left(\mathbf{E}\left[\widetilde{w}\right] - \rho_v\right) \approx v\left(\mathbf{E}\left[\widetilde{w}\right]\right) - \rho_v.v'\left(\mathbf{E}\left[\widetilde{w}\right]\right) \tag{6}$$

En égalisant les termes de droite de (5) et de (6), on a:

$$\frac{1}{2}v''\left(\mathbf{E}\left[\widetilde{w}\right]\right).\boldsymbol{\sigma}_w^2 = -\rho_v.v'\left(\mathbf{E}\left[\widetilde{w}\right]\right) \Rightarrow \rho_v = -\frac{1}{2}\frac{v''\left(\mathbf{E}\left(\widetilde{w}\right)\right)}{v'\left(\mathbf{E}\left(\widetilde{w}\right)\right)}$$

Comme on a l'expression symétrique pour toute autre fonction u, on retrouve analytiquement le lien entre courbure et prime de risque:

$$-\frac{u''}{u'} > -\frac{v''}{v'} \Leftrightarrow \rho_u > \rho_v$$

Lorsque les préférences vérifient l'utilité espérée, on peut donc également proposer comme mesure d'aversion à l'égard du risque le terme $-\frac{u''}{u'}$ pour toute fonction d'utilité élémentaire. Cette mesure constitue dans la littérature la mesure d'*Arrow-Pratt*.

[20] L'utilité élémentaire v est approximativement égale à:

$$v\left(w\right) \approx v\left(\mathbf{E}\left[\widetilde{w}\right]\right) + v'\left(\mathbf{E}\left[\widetilde{w}\right]\right).\left(w - \mathbf{E}\left[\widetilde{w}\right]\right) + \frac{1}{2}v''\left(\mathbf{E}\left[\widetilde{w}\right]\right).\left(w - \mathbf{E}\left[\widetilde{w}\right]\right)^2$$

L'expression approchée de l'utilité espérée est:

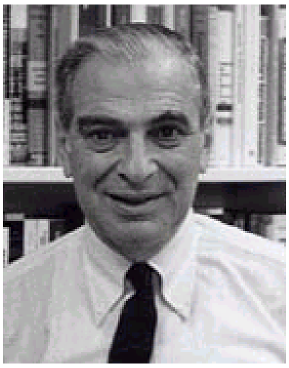
$$\begin{aligned} \mathbf{E}\left[v\left(w\right)\right] &\approx \mathbf{E}\left[v\left(\mathbf{E}\left[\widetilde{w}\right]\right) + v'\left(\mathbf{E}\left[\widetilde{w}\right]\right).\left(w - \mathbf{E}\left[\widetilde{w}\right]\right) + \frac{1}{2}v''\left(\mathbf{E}\left[\widetilde{w}\right]\right).\left(w - \mathbf{E}\left[\widetilde{w}\right]\right)^2\right] \\ &= \mathbf{E}\left[v\left(\mathbf{E}\left[\widetilde{w}\right]\right)\right] + \mathbf{E}\left[v'\left(\mathbf{E}\left[\widetilde{w}\right]\right).\left(w - \mathbf{E}\left[\widetilde{w}\right]\right)\right] + \mathbf{E}\left[\frac{1}{2}v''\left(\mathbf{E}\left[\widetilde{w}\right]\right).\left(w - \mathbf{E}\left[\widetilde{w}\right]\right)^2\right] \end{aligned}$$

Comme $v\left(\mathbf{E}\left[\widetilde{w}\right]\right)$, $v'\left(\mathbf{E}\left[\widetilde{w}\right]\right)$, $\frac{1}{2}v''\left(\mathbf{E}\left[\widetilde{w}\right]\right)$ sont des termes constants, on peut les sortir de l'opérateur espérance:

$$\mathbf{E}\left[v\left(w\right)\right] \approx v\left(\mathbf{E}\left[\widetilde{w}\right]\right) + v'\left(\mathbf{E}\left[\widetilde{w}\right]\right).\mathbf{E}\left[w - \mathbf{E}\left[\widetilde{w}\right]\right] + \frac{1}{2}v''\left(\mathbf{E}\left[\widetilde{w}\right]\right).\mathbf{E}\left[\left(w - \mathbf{E}\left[\widetilde{w}\right]\right)^2\right]$$

Evidemment $\mathbf{E}\left[w - \mathbf{E}\left[\widetilde{w}\right]\right] = \mathbf{E}\left[w\right] - \mathbf{E}\left[\widetilde{w}\right] = 0$, $\mathbf{E}\left[\left(w - \mathbf{E}\left[\widetilde{w}\right]\right)^2\right] = \sigma_w^2$ (= la variance de la richesse) et donc:

$$\mathbf{E}\left[v\left(w\right)\right] \approx v\left(\mathbf{E}\left[\widetilde{w}\right]\right) + \frac{1}{2}v''\left(\mathbf{E}\left[\widetilde{w}\right]\right).\boldsymbol{\sigma}_w^2$$



Figure~44: Kenneth J. Arrow

4.4 Equivalence des mesures

La mesure d'Arrow-Pratt n'est pas fondamentalement différente de l'approche de Yaari. En effet, on peut montrer que $-\frac{u''}{u'}$ non seulement mesure la courbure des fonctions mais également celle de la courbe d'indifférence au point w.

Pour voir ceci, écrivons l'équation implicite définissant cette courbe d'indifférence.

Notons :

- x_1 et x_2 les revenus supplémentaires dans l'état s_1 et l'état s_2;
- $x_2(x_1)$ la fonction résumant la courbe d'indifférence passant par le point w.

Cette fonction est alors implicitement donnée par l'équation suivant:

$$\pi_1 u(w + x_1) + \pi_2 u(w + x_2(x_1)) = u(w) \tag{7}$$

Si on différentie (7) par rapport à x_1, on obtient:

$$\pi_1.u'(w + x_1).dx_1 + \pi_2.u'\left(w + x_2(x_1)\right) x_2'(x_1).dx_1 = 0$$

d'où:

$$\pi_1.u'(w + x_1) + \pi_2.u'\left(w + x_2(x_1)\right) x_2'(x_1) = 0 \tag{8}$$

et donc la pente de la courbe d'indifférence est en x_1:

$$x_2'(x_1) = -\frac{\pi_1}{\pi_2}.\frac{u'(w + x_1)}{u'\left(w + x_2(x_1)\right)}$$

Par conséquent, au point $x_1 = 0$, cette pente est:

$$x_2^{'}(0) = -\frac{\pi_1}{\pi_2} \tag{9}$$

Ce résultat établit une propriété importante de l'utilité espérée: au voisinage de la certitude, un agent est approximativement neutre au risque et donc ses taux de substitution sont alors approximativement déterminés par les probabilités.

Une différentiation supplémentaire de la relation (8) permet alors de caractériser la courbure à l'origine. Comme:

$$\pi_1.u''(w+x_1).dx_1+\pi_2.[u''\left(w+x_2(x_1)\right)x_2^{'}(x_1)\times x_2^{'}(x_1)+u''\left(w+x_2(x_1)\right)x_2^{''}(x_1)]dx_1 = 0$$

on a après simplification:

$$\pi_1.u''(w+x_1)+\pi_2.u''\left(w+x_2(x_1)\right)\left(x_2^{'}(x_1)\right)^2+\pi_2 u'\left(w+x_2(x_1)\right)x_2^{''}(x_1) = 0 \tag{10}$$

Evaluée au point $x_1 = 0$, prenant en compte le résultat (9): $x_2'(0) = -\pi_1/\pi_2$, cette équation se réécrit:

$$\pi_1.u''(w)+\pi_2.u''(w)\left[-\frac{\pi_1}{\pi_2}\right]^2+\pi_2 u'(w)\,x_2^{''}(x_1) = 0$$

ou encore:

$$u''(w).\frac{\pi_1}{\pi_2}+\pi_2 u'(w)\,x_2^{''}(x_1) = 0$$

La courbure de la courbe d'indifférence à l'origine, $x_2^{''}(0)$, est donc:

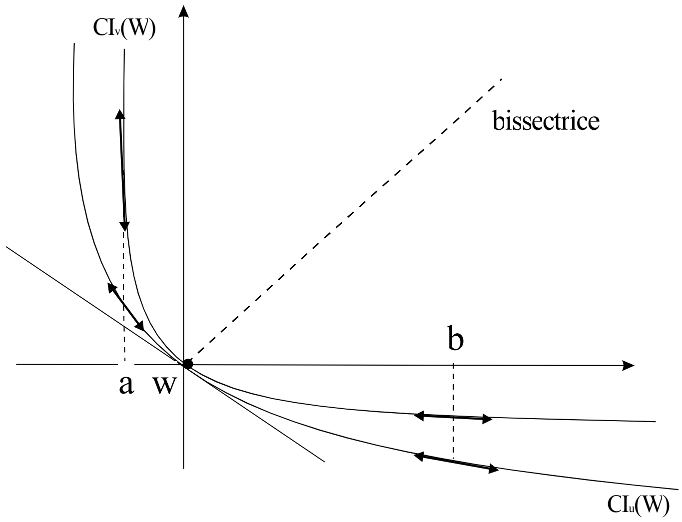
$$x_2^{''}(x_1) = \frac{\pi_1}{\pi_2^2}\left(-\frac{u''(w)}{u'(w)}\right) \tag{11}$$

Le terme π_1/π_2 étant une constante, la *mesure d'Arrow-Pratt* d'aversion au risque $-\frac{u''}{u'}$ correspond donc bien également (à un multiple près) à la courbure de la courbe d'indifférence en 0. Comme l'illustre le graphique 45, la mesure d'Arrow-Pratt détermine les inclusions respectives des positions acceptables:

$$-\frac{u''}{u'} > -\frac{v''}{v'} \Rightarrow AR_u > AR_v$$

En effet, si l'on note x_2^u et x_2^v sont les fonctions décrivant les courbes d'indifférence, comme à l'origine les pentes sont les mêmes:

$$\frac{dx_2^u}{dx_1}(0) = \frac{dx_2^v}{dx_1}(0) = -\frac{\pi_1}{\pi_2}$$



Figure~45: Courbure et taille de l'ensemble des positions acceptables

la décroissance plus rapide de u, i.e. la relation $-\frac{u''}{u'} > -\frac{v''}{v'}$ implique qu'à droite du point:

$$0 > \frac{dx_2^u}{dx_1}(0) > \frac{dx_2^v}{dx_1}(0) > -\frac{\pi_1}{\pi_2}$$

et donc:

$$0 > x_2^u(x_1) > x_2^v(x_1) \;\; \forall x_1 > 0$$

comme l'illustre le point b sur la figure 45.

A l'inverse à gauche de l'origine, partant de

$$\frac{dx_2^u}{dx_1}(0) = \frac{dx_2^v}{dx_1}(0) = -\frac{\pi_1}{\pi_2}$$

la décroissance plus rapide de u, i.e. la relation $-\frac{u''}{u'} > -\frac{v''}{v'}$, implique qu'à droite du point:

$$-\frac{\pi_1}{\pi_2} > \frac{dx^v}{dx_1}(0) > \frac{dx_2^u}{dx_1}(0)$$

et donc:

$$x_2^u(x_1) > x_2^v(x_1) \;\; \forall x_1 < 0$$

comme l'illustre le point a sur la figure 45. Nécessairement donc, on voit bien que les deux approches de l'aversion à l'égard du risque se confondent dès lors que l'hypothèse d'utilité espérée est admise:

$$AR_u > AR_v \Leftrightarrow -\frac{u''}{u'} > -\frac{v''}{v'}$$

En fait, on peut même élargir cette équivalence à la troisième mesure de l'aversion, la prime de risque. Ceci constitue le *théorème de Pratt* [Pra64] dont la démonstration est en annexe:

Théorème 2 *Soient u et v deux fonctions d'utilité élémentaires définies sur la richesse; les trois propriétés suivantes sont équivalentes :*
(i) $-\frac{u''}{u'} > -\frac{v''}{v'}$ *pour tout w;*
(ii) il existe une fonction croissance strictement concave G telle que:

$$u(w) = G(v(w))$$

(iii) si la richesse s'écrit $w + \widetilde{\epsilon}$, où w est une richesse certaine, $\widetilde{\epsilon}$ est un bruit blanc - $\mathbf{E}\widetilde{\epsilon} = 0$ *- alors:*

$$\rho_u(w) > \rho_v(w)$$

4.5 Demande d'assurance

L'utilité espérée est une représentation des préférences obtenue au prix d'hypothèses supplémentaires (trop fortes?). La contrepartie de ce coût sont des résultats analytiques plus nombreux. L'un des plus remarquables concerne la demande d'assurance.

Un agent dispose d'une richesse initiale W et peut perdre avec une probabilité p une somme L. Pour se protégér de cette éventualité, l'agent peut recourir à une assurance lui donnant une indemnité q en cas de dommage; le prix de l'assurance est πq, π étant le taux d'assurance.
Le problème d'optimisation de l'agent s'il vérifie l'hypothèse d'utilité espérée est donc simplement:

$$\max_q \; pu(W - L - \pi q + q) + (1 - p)u(W - \pi q)$$

La condition de premier ordre (cpo) de ce problème est :

$$pu'(W - L - \pi q + q)(1 - \pi) - (1 - p)u(W - \pi q)\pi = 0$$

Lorsque les conditions de stricte concavité sur u sont vérifiées, la cpo est suffisante et donc le choix optimal q^* est définie par:

$$\frac{u'(W - L + (1 + \pi)q^*)}{u'(W - \pi q^*)} = \frac{\pi}{1 - \pi}\frac{1 - p}{p} \qquad (12)$$

Une assurance complète est une situation où l'agent a des revenus stables:

$$W - L + (1 + \pi)q = W - \pi q$$

et donc où:

$$q = L$$

En une telle situation le rapport des utilités marginales est:

$$\frac{u'(W - L + (1+\pi)q)}{u'(W - \pi q)} = \frac{u'(W - \pi q)}{u'(W - \pi q)} = 1 \tag{13}$$

L'assurance parfaite ne sera donc la solution optimale, c'est-à-dire la solution de (12) que si:

$$\frac{\pi}{1-\pi}\frac{1-p}{p} = 1$$

Ce résultat permet donc de déterminer la demande d'assurance en fonction de la position du taux d'assurance π par rapport à la probabilité d'occurence de la perte p. Si $\pi > p$ alors $\frac{\pi}{1-\pi}\frac{1-p}{p} > 1$ et donc on doit avoir:

$$\frac{u'(W - L + (1+\pi)q)}{u'(W - \pi q)} > 1 \Rightarrow u'(W - L + (1+\pi)q) > u'(W - \pi q)$$

Comme l'utilité marginale est décroissant, le choix optimal conduit à:

$$W - L + (1+\pi)q < W - \pi q$$

c'est-à-dire à une assurance incomplète:

$$q < L$$

Si $\pi < p$, les relations inverses conduisent évidemment à une sur-assurance:

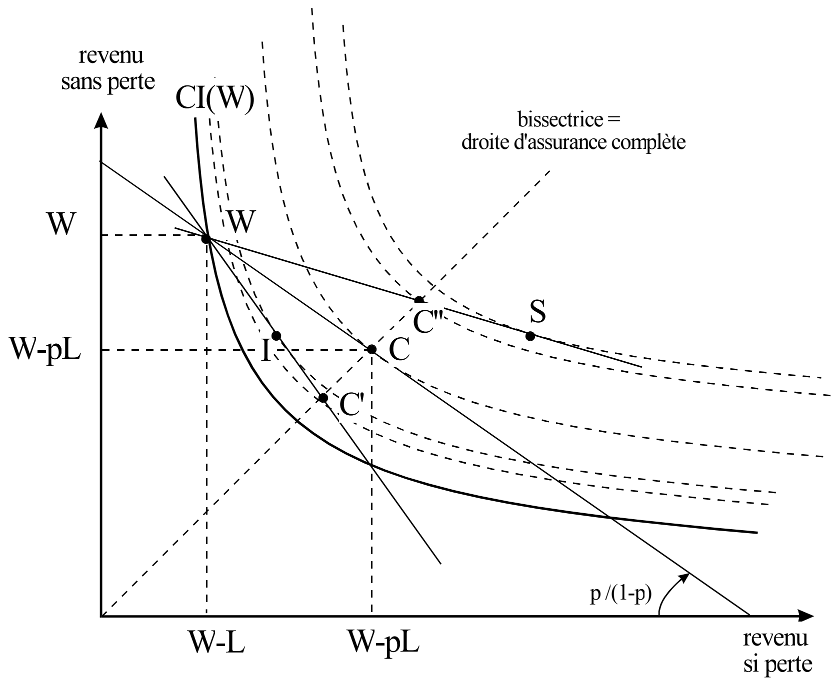
$$q > L$$

Dans cette situation, l'agent aura alors un revenu supérieur lorsque la perte se réalise!

Si l'on suppose que les compagnies d'assurance sont capables d'annuler le risque en le mutualisant, en faisant jouer la Loi des Grands Nombres, le profit moyen effectif sur chaque police sera alors $\pi q - p.q$ puisque la compagnie perçoit *toujours* la prime d'assurance, πq, mais doit, avec une probabilité p, verser l'indemnité q. Ce profit est linéaire (= rendement constant). Aussi, si la concurrence est parfaite, le profit à l'équilibre sera nul. Ceci n'est possible que si:

$$\pi q - p.q = 0 \Rightarrow \pi = p$$

A l'équilibre, les agents seront donc complètement assurés si les risques sont mutualisables.



Figure~46: La demande d'assurance pour des préférences strictement convexes

Ce résultat est sans doute l'un des plus intéressants de l'utilité espérée. Comme l'illustre la figure 46, ce résultat peut cependant être généralisé à des préférences strictement convexes dont les taux marginaux de substitution sont égaux aux rapports des probabilités[21]:

$$Tms_{1\to 2}(w,w) = \frac{p}{1-p} \ \forall w > 0$$

En utilisant l'hypothèse de stricte convexité des préférences (et ses conséquences usuelles) sur les Tms, on retrouve les résultats précédents:

- si l'assurance est équitable $\pi = p$, la contrainte budgétaire est la droite passant par les points W et C; le choix optimal, C, est l'assurance complète;
- si le coût de l'assurance est élevé, $\pi > p$, la droite budgétaire est alors la droite passant par les points W et I; comme $\forall w > 0\ Tms_{1\to 2}(w,w) =$

[21] Et on porrait encore relâcher ce résultat en abandonnant l'hypothèse de dérivabilité des fonctions, l'existence d'un Tms, pour d'autres hypothèses plus faibles.

$\frac{p}{1-p} < \frac{\pi}{1-\pi}$, partant du point W, l'agent ne va pas sur cette droite budgétaire jusqu'au point C' d'assurance complète trop coûteux; le choix optimal, I, est donc un point au dessus de la bissectrice et l'assurance est donc incomplète;

- si le coût de l'assurance est faible, $\pi < p$, la droite budgétaire est alors la droite passant par les points W et S; comme $\forall w > 0\ Tms_{1\to 2}(w,w) = \frac{p}{1-p} > \frac{\pi}{1-\pi}$, partant de W, l'agent ne s'arrête pas au point C'' d'assurance complète en raison du coût très faible de l'assurance; son choix optimal, S, est un point au dessous de la bissectrice et on a sur-assurance.

4.6 Aversion relative à l'égard du risque

Les différentes notions d'aversion à l'égard du risque sont définies dans l'espace des revenus supplémentaires *partant d'une richesse initiale certaine.* Intuitivement, on cherche donc à déterminer pour chaque agent les profils de revenu supplémentaire pour lesquels il accepterait de supporter les risques supplémentaires proposés. Comme ces revenus peuvent être le produit d'investissements, d'actifs financiers, on aurait également pu s'intéresser aux rendements nécessaires pour inciter l'agent à supporter des risques. Ceci revient en fait tout simplement à redéfinir nos notions dans l'espace des rendements (ou des rendements nets) de ces positions risquées.

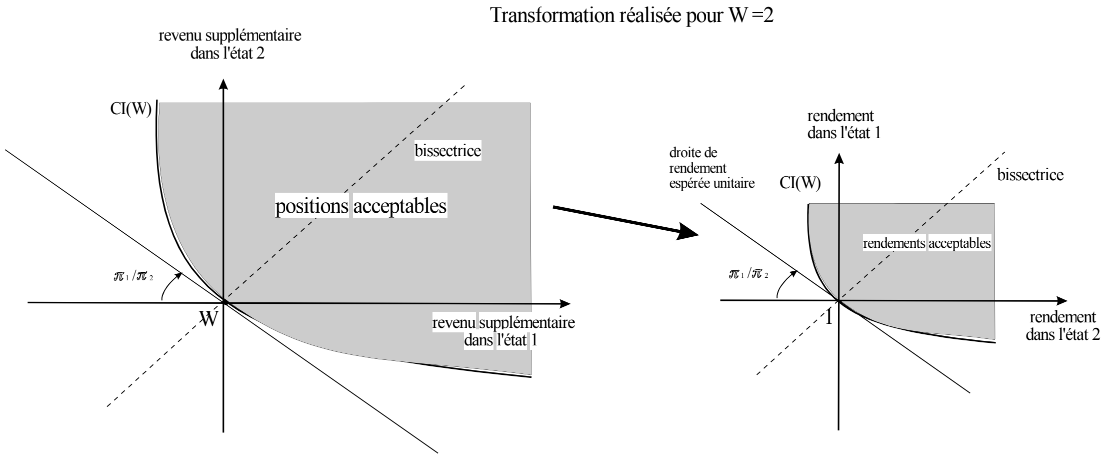
Formellement, en supposant toujours que $\#S = 2$, si l'on note U la fonction d'utilité, w la richesse initiale, R_1 et R_2 les rendements nets. La richesse finale s'écrit donc $R_s.w$ avec $s = 1,2$. Si l'on cherche donc à déterminer les *rendements acceptables* pour l'agent, on définit alors l'ensemble $R_U(w)$:

$$R_U(w) = \left\{(R_1, R_2) \in \Re^2_+ : U(R_1.w, R_2.w; \pi_1, \pi_2) \geq U(w, w; \pi_1, \pi_2)\right\}$$

Cet ensemble des rendements acceptables est évidemment en relation biunivoque avec l'ensemble des positions acceptables: toute position acceptable (z_1, z_2) de $A_U(w)$ définit évidemment un vecteur de rendements aceptable:

$$(z_1, z_2) \in A_U(w) \Rightarrow \left(\frac{z_1}{w}, \frac{z_2}{w}\right) \in R_U(w)$$

Graphiquement, l'ensemble $R_U(w)$ est donc obtenue de $A_U(w)$ par une homothétie de rapport $1/w$ et de centre (w,w). Aussi, lorsque les préférences sont strictement convexes, elle est également un ensemble convexe, compris dans un demi-plan - figure 47.



Figure~47: Détermination de l'ensemble des rendements acceptables(à droite) à partir des positions acceptables (à gauche)

La prime de risque, ρ_U^R, est redéfinie comme la perte maximale de rendement que l'individu est prêt à consentir pour obtenir un rendement certain:
ρ_U^R / $U\left(\left[\mathbf{E}\left[\widetilde{R}\right]-\rho_U^R\right].w,\left[\mathbf{E}\left[\widetilde{R}\right]-\rho_U^R\right].w;\pi_1,\pi_2\right)=U\left(R_1.w,R_2.w;\pi_1,\pi_2\right)$

Lorsque l'utilité vérifie l'hypothèse d'utilité espérée, on peut également proposée une mesure d'aversion similaire à la mesure d'Arrow-Pratt. Pour la déterminer on recourt aux développements limités autour du rendement initial $R=1$:

$$u\left(R.w\right)\approx u\left(w\right)+u'\left(w\right).\left(R-1\right).w+\frac{1}{2}u''\left(w\right).\left[\left(R-1\right)w\right]^2$$
$$u\left(w\left(1-\rho_u^R\right)\right)\approx u\left(w\right)-\rho_u^R w.u'\left(w\right)$$

Par conséquent, par des calculs similaires à ceux déjà vus (page 51) :

$$\mathbf{E}\left[u\left(\widetilde{R}.w\right)\right]\approx u\left(w\right)+\frac{1}{2}u''\left(w\right).\boldsymbol{\sigma}_R^2.w^2$$

où $\boldsymbol{\sigma}_R^2=\mathbf{E}\left[\left(\widetilde{R}-1\right)\right]$ est la variance du rendement. Aussi, en égalisant les approximations de $u\left(w\left(1-\rho_u^R\right)\right)$ et de $\mathbf{E}\left[u\left(\widetilde{R}.w\right)\right]$, on obtient:

$$-\rho_u^R w.u'\left(w\right)=\frac{1}{2}u''\left(w\right).\boldsymbol{\sigma}_R^2.w^2\Rightarrow\rho_u^R=\frac{1}{2}\left(-\frac{w.u''\left(w\right)}{u'\left(w\right)}\right)\sigma_R^2$$

La prime de risque (exprimée en rendement) est donc approximativement proportionnelle au terme $-\frac{w.u''}{u'}$: ce terme constitue donc l'*aversion relative au risque* - la mesure $-\frac{u''}{u'}$ étant qualifiée d'*aversion absolue au risque.*

5 Dérivation de l'utilité espérée

Cette section expose la démonstration du théorème de l'utilité espérée.

5.1 L'axiome d'indépendance

Pour obtenir cette restriction supplémentaire, des hypothèses supplémentaires sont introduites. Parmi les restrictions les plus fréquentes figurent les deux suivantes[22]:

Axiome 2 (Archimedien) *Soient trois loteries x, y, $z \in £$ vérifiant:*

$$x \succ y \succ z$$

alors il existe deux réels α, $\beta \in]0, 1[$ pour lesquels:

$$\alpha * x \oplus (1-\alpha) * z \succ y \succ \beta * x \oplus (1-\beta) * z$$

Axiome 3 (independance) *Soient deux loteries x, $y \in £$ vérifiant:*

$$x \succ y$$

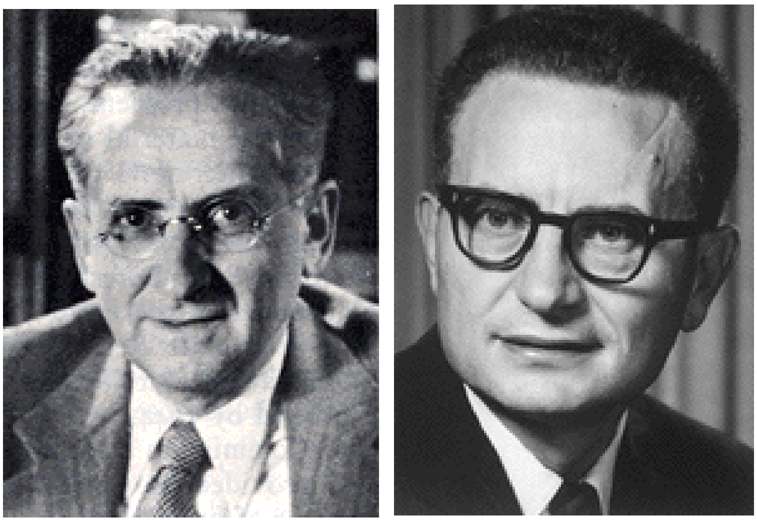
alors pour toute loterie $z \in £$, pour tout réel $\alpha \in]0, 1[$ on a:

$$\alpha * x \oplus (1-\alpha) * z \succ \alpha * y \oplus (1-\alpha) * z$$

L'axiome archimédien a un parfum d'hypothèse de continuité et n'est guère exigeant. La seconde hypothèse est moins technique et beaucoup plus exigeante. Introduite indépendamment par Marschak [Mar50] et Samuelson [Sam52], [23] elle fut très discutée et fréquemment invalidée par les expériences. Cette hypothèse revient en fait à exiger que si l'on préfère strictement une loterie à une autre, tout enrichissement symétrique des loteries ne viendra pas modifier nos préférences sur les loteries. Ainsi, sur la figure 49, on a le choix entre, d'une part, une loterie a_1 donnant 100 unités de numéraire avec une probabilité de 0.2, rien sinon et, d'autre part, une loterie a_2 donnant seulement un prix de 50 unités mais avec une probabilité de 0.3. Si l'on

[22] Ces deux hypothèses ne sont pas celles de von Neumann et Morgenstern de 1944. Cette présentation fut introduite par l'économiste N. Jensen seulement en 1967 [Jen67]. Elle est cependant l'une des plus courantes dans les *text books*; ainsi, elle est celle adoptée par Kreps dans son manuel [Kre90] auquel le lecteur pourra se référer pour plus de développements ainsi qu'au livre du même auteur [Kre88].

[23] Mais si elle n'était pas une hypothèse de von Neumann & Morgenstern, elle était une conséquence de leurs différents axiomes comme le démontra Edmond Malinvaud [Mal52].



Figure~48: Jacob Marschak et Paul Samuelson

préfère a_1 à a_2 [24], l'axiome d'indépendance implique alors que si on l'enrichit identiquement ces deux premières loteries pour obtenir les loteries b_1 et b_2 alors les choix doivent conserver une certaine cohérence:

$$a_1 \succ a_2 \Rightarrow b_1 \succ b_2$$

puisque les deux dernières loteries ne sont que des loteries composées obtenues à l'aide d'une loterie δ_{10} donnant 10 avec certitude:

$$b_1 = 0.9 * a_1 \oplus 0.1 * \delta_{10}$$

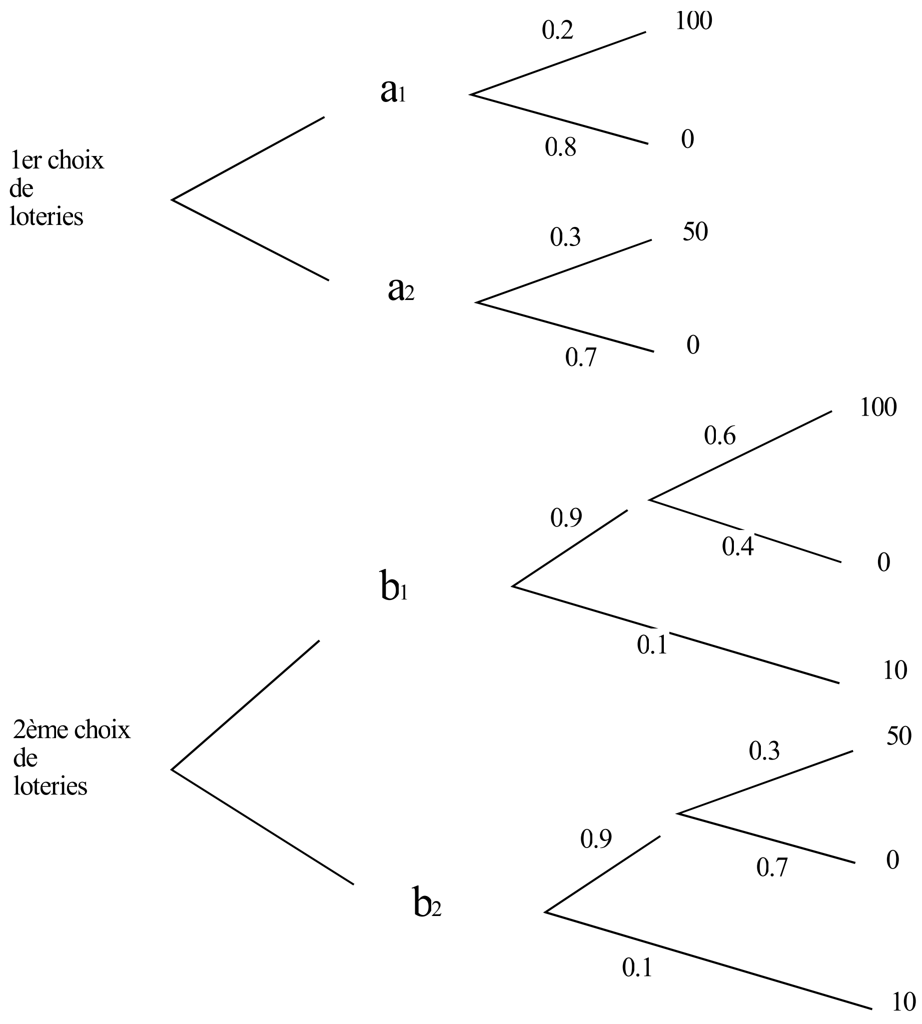
$$b_2 = 0.9 * a_2 \oplus 0.1 * \delta_{10}$$

A beaucoup sans doute, l'axiome d'indépendance dans le cas de la figure 49 n'apparaîtra pas comme très exigeant. Celui de la figure 50 mettra sans doute plus en difficulté la plausibilité de l'axiome d'indépendance. Dans cet exemple, on offre tout d'abord le choix entre les loteries suivantes:

- la loterie a_1, contre une probabilité de 0.4 de ne rien gagner, permet de gagner 100 avec une probabilité de 0.6;

- la loterie a_2, contre une probabilité de 0.36 de ne rien gagner, ne permet que de gagner 100 avec une probabilité de 0.54, de gagner 50 avec une probabilité de 0.1.

[24] Cependant une préférence pour a_2 ne doit pas être considérée comme un acte honteux. Tout est affaire de goût: "de gustibus non est disputandum".



Figure˜49: Axiome d'indépendance et choix de loteries, I

Le second choix de loteries donne le choix entre les loteries suivantes:

- la loterie b_1, contre une probabilité de 0.4 de ne rien gagner, permet de gagner 100 avec une probabilité de 0.6;
- la loterie b_2, avec un risque nul de rien gagner, assure un gain de 50.

Dans ce second choix, beaucoup préfèreront respecter une version adaptée d' "un tien vaut mieux que deux tu l'auras" et préfèreront la loterie b_2 plutôt que de courir le risque de ne rien obtenir avec une probabilité de 0,4. Dans le premier choix, sans doute également nombreux seront ceux qui préféreront la première loterie à la seconde car, pour une augmententation de seulement 0.04 de la probabilité de ne rien gagner, elle permet décrocher des prix de 100. Ces choix sont cependant incohérents. En effet, les loteries a_1 et a_2 sont des loteries composées obtenues à partir des deux autres:

$$a_1 = 0.9 * b_1 + 0.1 * b_1$$

$$a_2 = 0.9 * b_1 + 0.1 * b_2$$

Par conséquent, si l'axiome d'indépendance est respectée nécessairement les seuls choix permis sont les suivants:

$$a_1 \succ a_2 \Leftrightarrow b_1 \succ b_2$$

$$a_2 \succ a_1 \Leftrightarrow b_2 \succ b_1$$

L'hypothèse d'indépendance n'est donc pas une hypothèse mince. Aussi, durant les 20 dernières années, de nombreux travaux ont été réalisés pour relâcher cette hypothèse en sauvant l'utilité espérée.[25] Voyons cependant les résultats que cette hypothèse coûteuse permet d'obtenir.

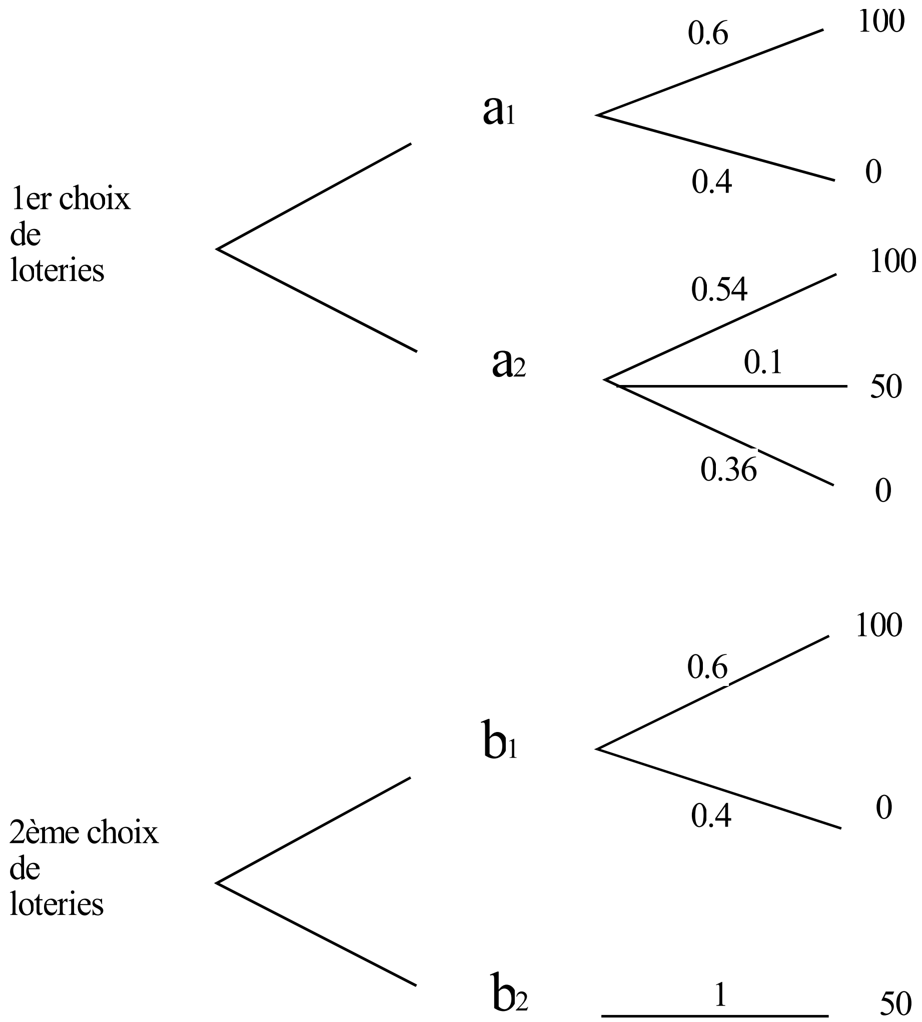
5.2 La propriété d'utilité espérée

Les préférences sont supposées représentées par un pré-ordre complet vérifiant les axiomes 2 et 3. Pour simplifier l'analyse (mais sans perte de généralité), les deux hypothèses suivantes sont introduites:

Hypothèse 4 *Le support de distribution des probabilités X est supposé être un ensemble fini.*

Hypothèse 5 *L'ensemble des loteries est 'borné': $\exists\ b, w \in \pounds$ t.q. $\forall x \in \pounds,\ w \preceq x \preceq b$.*

[25] Le lecteur intéressé pourra consulter notamment [Kre88], [Mac87].



Figure~50: Axiome d'indépendance et choix de loterie, II

Il est possible de montrer que les axiomes - et notamment l'axiome d'indépendance - imposent aux préférences une structure remarquable. Celle-ci est résumée par trois lemmes intermédiaires et un théorème de représentation.

Lemme 1 *Pour tout α, $\beta \in [0,1]$,*

$$\alpha * b \oplus (1-\alpha) * w \succ \beta * b \oplus (1-\beta) * w \Leftrightarrow \alpha > \beta$$

Démonstration: Comme $\forall \lambda \in]0,1[$:

$$b \succ \lambda * b \oplus (1-\lambda) * w$$

l'axiome d'indépendance assure en effet que :

$$\begin{aligned} \alpha * b \oplus (1-\alpha) * w &\succ \alpha * (\lambda * b \oplus (1-\lambda) * w) \oplus (1-\alpha) * w \\ &= \alpha\lambda * b \oplus (1-\alpha\lambda) * w \end{aligned}$$

Comme $\lambda \in]0,1[$:

$$\alpha > \lambda\alpha$$

En notant $\beta = \lambda\alpha$, on vient de démontrer que pour tout (α, β) tel que $\alpha > \beta$, on a:

$$\alpha * b \oplus (1-\alpha) * w \succ \beta * b \oplus (1-\beta) * w$$

La réciproque est immédiate. ■

Ce résultat est parfois présenté comme un axiome: dans ce cas, ceci revient à supposer qu'il est toujours préférable dans les loteries composées d'augmenter le poids de la “bonne” loterie. Le second résultat intermédiaire est le suivant:

Lemme 2 $\forall x, y, z \in \pounds$ *telles que:*

$$x \succ y \succ z$$

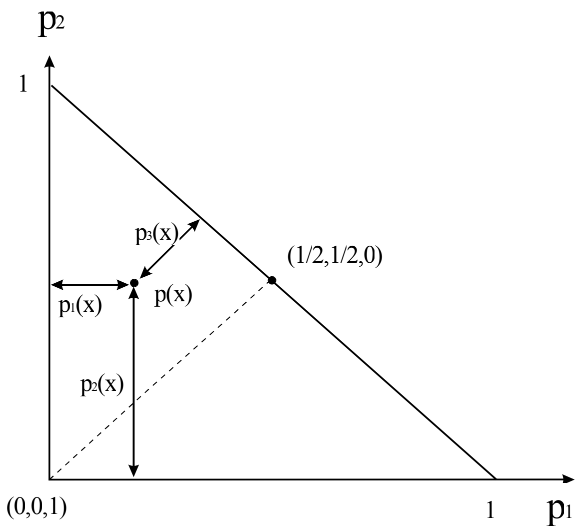
il existe un et un seul α vérifiant:

$$y \sim \alpha * x \oplus (1-\alpha) * z$$

Le corollaire immédiat de ce lemme est le suivant:

Corollaire 1 $\forall y \in \pounds$, *il existe un et un seul α vérifiant:*

$$y \sim \alpha * b \oplus (1-\alpha) * w$$

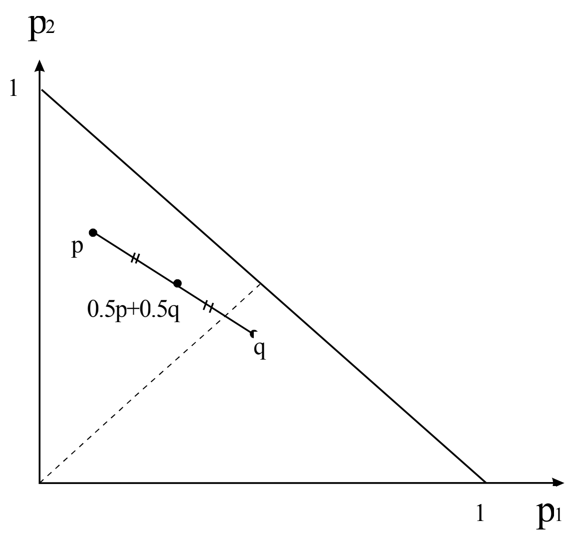


Figure~51: Le triangle de Marschak

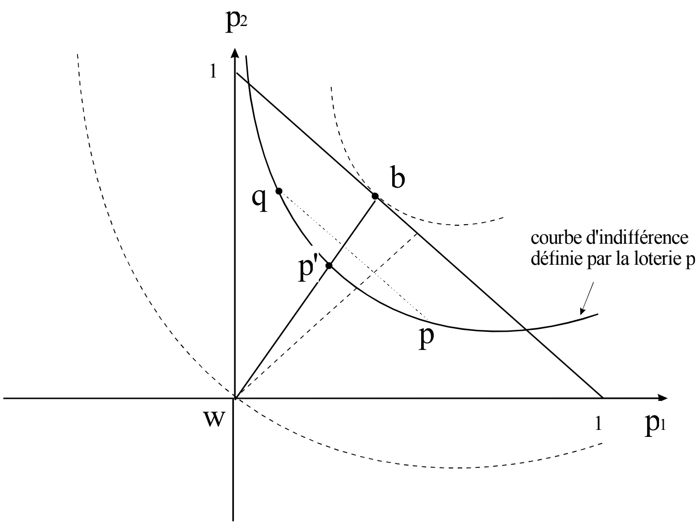
Ce corollaire est très important car il assure la possibilité d'adapter la construction de Wold à l'incertain: on peut en effet repérer la position de chaque courbe d'indifférence non par son intersection avec une droite dans l'espace des paniers (cf figure 20) mais par sa position sur le segment reliant les deux loteries b et w. Ainsi, si l'on suppose que le support des loteries X se réduit à trois valeurs $\{x_1, x_2, x_3\}$ les loteries peuvent être représentées dans le triangle de Marschak (1950). En effet, $\{x_1, x_2, x_3\}$ étant donné, chaque loterie x est alors uniquement définie par les probabilités de chaque prix $p_x = (p_1(x), p_2(x), p_3(x))$. Comme l'illustre la figure 51, si l'abscisse et l'ordonnée permettent de lire les deux premières coordonnées de chaque vecteur de probabilités p_x, la dernière est donnée par la distance du point p_x à la frontière du simplexe, i.e. à la droite reliant les points $(1, 0)$ et $(0, 1)$ d'équation: $p_2 = 1 - p_1$. Chaque point du triangle résume donc graphiquement une loterie. Les loteries composées sont des barycentres des loteries simples combinées. Ainsi, sur la figure 52, partant des deux loteries définies par les probabilités p et q, on a construit la loterie composée $0.5 * [(x_1, x_2, x_3), p] \oplus 0.5 * [(x_1, x_2, x_3), q]$: graphiquement, celle-ci est le milieu du segment $[p, q]$. Chaque loterie (résumée par ses probabilités) p définit une courbe d'indifférence $CI(p)$:

$$CI(p) = \{q \in \Delta(X) : [x, q] \sim [x, p]\}$$

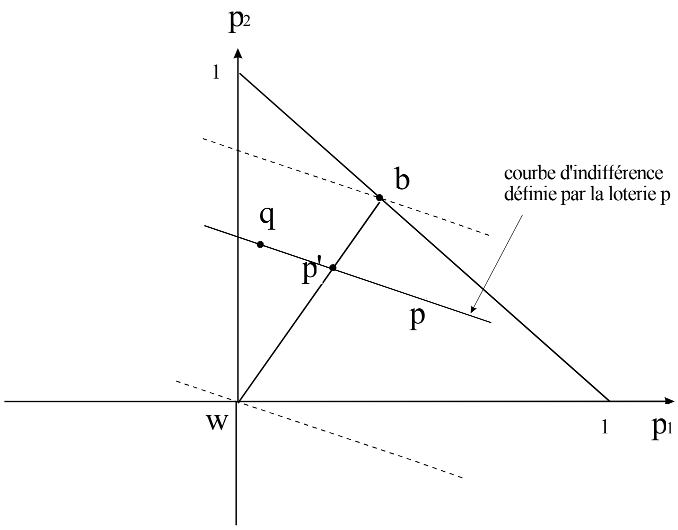
Un exemple (*arbitraire*) de telles courbes d'indifférence est représenté sur la figure 53. Comme dans la construction de Wold, une fois les courbes



Figure~52: Détermination graphique d'une loterie composée



Figure~53: Les courbes d'indifférence dans le triangle de Marschak: un exemple



Figure~54: Courbes d'indifférence respectant l'axiome d'indépendance (et l'axiome archimédien).

d'indifférence définies, une fois les deux loteries extrêmes b et w introduites, on pourra définir un indice d'utilité: par exemple, sur la figure 53, il suffira pour tout point p de définir le point de sa courbe d'indifférence coupant le segment $[w, b]$, le point p' sur le graphique; comme p' est une loterie composée de b et de w, puisqu'il est sur leur segment, il suffira pour définir un indice d'utilité de déterminer le poids $U(p) \in [0,1]$ pour lequel:

$$p \sim U([x,p]) * b \oplus (1 - U([x,p])) * w$$

En effet, en raison du lemme 1:

$$U([x,p]) > U([x,q]) \Leftrightarrow [x,p] \succ [x,q]$$

Le troisième résultat intermédiaire essentiel est le suivant:

Lemme 3 *Pour tout* $x, y \in \pounds$ *telles que:*

$$x \sim y$$

alors $\forall z \in \pounds$, $\forall \alpha \in]0,1[$:

$$\alpha * x \oplus (1-\alpha) * z \sim \alpha * y \oplus (1-\alpha) * z$$

Le résultat précédent est très important car de lui on en déduit une des propriétés les plus fortes des préférences respectant l'axiome d'indépendance: la linéarité des courbes d'indifférence. En effet, si pour tout couple loteries équivalentes x, y de les combiner à une d'elle, par exemple x, le lemme 3 nous assure que:

$$x \sim y \Rightarrow \forall \alpha \in]0,1[: x = \alpha * x \oplus (1-\alpha) * x \sim \alpha * y \oplus (1-\alpha) * x$$

Dans le triangle de Marschak, les combinaisons convexes de deux loteries équivalentes sont équivalentes. Aussi, les courbes d'indifférence ne peuvent être strictement convexes (ou concaves) comme sur la figure 53 (page 67). Nécessairement, les courbes d'indifférence sont des droites comme sur la figure **??** parallèles les unes aux autres. Cette propriété géométrique traduit en fait la propriété d'utilité espérée comme le théorème suivant le montre.

Théorème 3 *Sous les axiomes 3 et 2, les hypothèses 1, 2, 3, 5 et 4, pour tout pré-ordre complet $\succeq$ défini sur un ensemble £ il existe une fonction d'utilité U représentant ces préférences et satisfaisant la propriété d'espérance d'utilité:*

$$U\left(\left[(x_i)_{i=1,\ldots}, (p_x(x_i))_{i=1,\ldots}\right]\right) = \sum_i p_x(x_i) . u(x_i)$$

où u est la fonction d'utilité élémentaire de U.

démonstration : La démonstration est en annexe. Elle consiste à appliquer la démarche de Wold en définissant la fonction u de la manière suivante : $u(w) = 0$, $u(b) = 1$. La propriété de parallélisme impose (cf figure 54) alors que pour tout prix x il existe un unique réel $u(x) \in [0,1]$ pour lequel:

$$x \sim u(x) * b \oplus (1 - u(x)) * w$$

■

Le résultat peut être considérablement étendu. Ainsi, le théorème 3 peut être prolongé au cas où il n'existe ni de loterie b dominante, ni de loterie étendue. De même, si l'on suppose que les préférences sont continues, si l'ensemble X vérifie certaines propriétés topologiques habituelles pour les espaces euclidiens, ce résultat tient toujours lorsque X n'est plus nécessairement fini. Au surplus, sous les conditions mentionnées, la fonction d'utilité élémentaire est alors une fonction continue. Notons que cette fonction d'utilité élémentaire a une propriété remarquable: elle est une *mesure cardinale des choix*. En effet, lorsque par exemple $u(x) - u(y) = 2[u(z) - u(y)]$, ceci implique:

$$u(x) = \frac{1}{2}u(y) + \frac{1}{2}u(z)$$

La relation $u(x) - u(y) = 2[u(z) - u(y)]$ a donc un sens économique: pour l'agent considéré, recevoir la loterie $\left[(y,z), \left(\frac{1}{2}, \frac{1}{2}\right)\right]$ ou avec certitude x est indifférent.

Comme pour toute fonction d'utilité, on peut toujours transformer la fonction d'utilité U par une fonction monotone strictement croissante G; la fonction ainsi obtenue $V = G \circ U$ représente toujours les mêmes préférences:

$$G \circ U(a_x) > G \circ U(a_y) \Leftrightarrow a_x \succ a_y$$

Cependant l'information économique contenue dans la fonction d'utilité élémentaire est évidemment perdue. Pour préserver cette propriété de cardinalité seules des transformations affines sont possibles.

Cette propriété de cardinalité suscita bien des suspicions. En effet, la révolution parétienne des années 30, initiée notamment par Hicks et Allen, l'utilité cardinale était regardée comme une notion métaphysique, héritage des errements utilitaristes de l'économie. Aussi, l'utilité espérée fut *initialement* rejetée par certains économistes, et non des moindres: Paul A. Samuelson [Sam50], William J. Baumol [Bau51]. Le choix semblait être le suivant: "préférences ordinales ou utilité cardinale" pour reprendre le titre d'un article de Wold [Wol52].

Mais l'utilité espérée de von Neumann et Morgenstern n'est pas celle de Daniel Bernoulli. Chez ce dernier, l'utilité élémentaire u est d'abord définie pour obtenir *in fine* l'utilité espérée U. Dans l'approche axiomatique moderne, par contre, les données initiales sont les préférences (sur les loteries). De celles-ci est déduite une fonction d'utilité *ordinale* U et, seulement après, sous des hypothèses supplémentaires sur les préférences, celle-ci est *décomposée* pour obtenir l'utilité espérée. Par conséquent, l'utilité élémentaire à la von Neumann est une "utilité cardinale qui est ordinale" pour reprendre l'heureuse expression par laquelle William J. Baumol [Bau58] referma ce débat en 1958.

6 Risque et incertitude*

6.1 L'incertain statut des probabilités

Dans l'approche à la von Neumann-Morgenstern, l'analogie avec les loteries permet d'opérer dans un cadre assez général. Cependant, l'analogie ne suppose pas seulement que les résultats ou les conséquences soient aléatoires. Une loterie n'est pas en effet seulement caractérisée par des prix aléatoires mais aussi par une distribution *objective*: la probabilité que l'on obtienne ainsi un six, un double six au lancé de dés est une variable connues de tous

Figure~55: Pierre-Simon de Laplace et Richard von Mises

les joueurs, commune à ceux-ci, mesurable. La subjectivité de l'acteur n'a aucune part dans sa détermination. Ceci s'accorde avec une certaine vision "classique" de certains mathématiciens influencés notamment par les jeux de hasard ou par des recherches statistiques (Bernoulli, de Moivre, Gauss, Laplace). Ainsi, Laplace (1749-1827) présentait ainsi en 1812 son ouvrage *Théorie analytique des probabilités*[26]:

> "Mon objet étant de présenter ici les méthodes et les résultats généraux de la théorie des probabilités, je traite spécialement les questions les plus délicates, les plus difficiles, et en même temps les plus utiles de cette théorie. Je m'attache surtout, à déterminer la probabilité des causes et des résultats indiqués par les événements considérés en grand nombre, et à chercher les lois suivant lesquelles cette probabilité approche de ses limites, à mesure que les événements se multiplient." ([Lap12] p.2)

Ceci ne conduisait pas nécessairement ces auteurs à croire en l'existence du hasard, en sa réalité ontologique:

> "tous les événements, ceux même qui par leur petitesse, semblent ne pas tenir aux grandes lois de la nature, en sont une suite

[26] Cet ouvrage reprend en fait les conceptions déjà exposées par Laplace en 1795 dans son *Essai philosophique sur les probabilités* [Lap95].

aussi nécessaire que les révolutions célestes. Une intelligence qui pour un instant donné, connaîtrait toutes les forces dont la matière est animée, ainsi que la position et la vîtesse de chacune de ses molécules ; si d' ailleurs elle était assez vaste pour soumettre ces données à l' analyse, embrasserait dans la même formule, les mouvements les plus grands corps de l' univers et ceux du plus léger atome. Pour une semblable intelligence, rien ne serait irrégulier, et la courbe décrite par une simple molécule d' air ou de vapeurs, paraîtrait réglée d' une manière aussi certaine, que l' est pour nous l' orbe du soleil. Mais dans l'ignorance où nous sommes de l'immensité des données nécessaires à la solution de ce grand problème, et dans l'impossibilité, vu notre faiblesse, d'assujetir au calcul la plupart de celles qui nous sont connues, alors même que leur nombre est très borné ; nous attribuons les phénomènes qui nous paraissent arriver et se succéder sans aucun ordre, à des causes variables et cachées, dont l' action a été désignée par le mot hasard, mot qui n'est au fond que l'expression de notre ignorance. La probabilité est relative, en partie, à cette ignorance, et en partie, à nos connaissances. Nous savons que sur trois ou un plus grand nombre d' événemens, un seul doit exister ; mais rien ne porte à croire que l' un d' eux arrivera plutôt que les autres. Dans cet état d'indécision, il est impossible de prononcer avec certitude sur leur existence. Il est cependant probable qu'un de ces événements, pris à volonté, n'existera pas; parce que, sur divers cas pour nous également possibles, nous en voyons plusieurs qui excluent son existence, tandis qu'un seul la favorise. La théorie des probabilités consiste à réduire tous les événements qui peuvent avoir lieu dans une circonstance donnée, à un certain nombre de cas également possibles, c' est-à-dire tels que nous soyons également indécis sur leur existence, et à déterminer parmi ces cas, le nombre de ceux qui sont favorables à l' événement dont on cherche la probabilité. Le rapport de ce nombre à celui de tous les cas possibles, est la mesure de cette probabilité qui n'est donc qu' une fraction dont le numérateur est le nombre des cas favorables, et dont le dénominateur est celui de tous les cas possibles."
([Lap12] pp. 178-79)

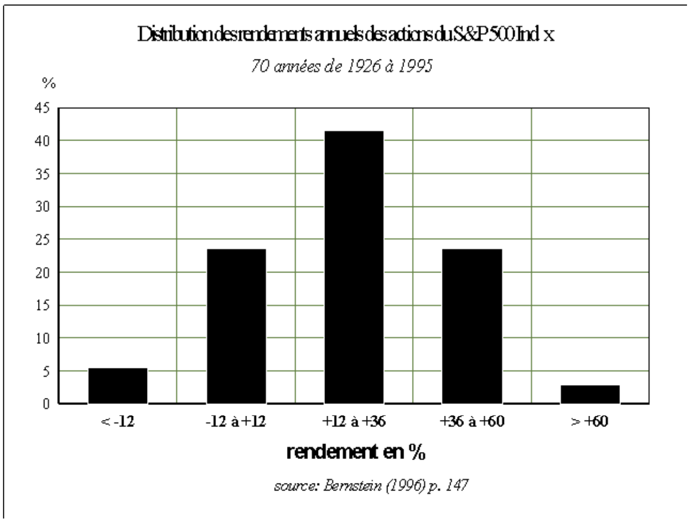
Au XXème siècle, cette conception classique a été réincarnée notamment par les partisans de l'approche "fréquentiste" exposée notamment par le statisticien Richard von Mises (1883-1953) [Mis54], un des pères des mathématiques appliquées[27]. La probabilité d'un événement y définie comme la valeur limite vers laquelle doit tendre cette événement lorsque l'expérience qui l'engendre est répétée à l'infini. L'ombre de Jacob Bernoulli se profile donc derrière cette conception positive: il n'est pas de probabilité en dehors du domaine de la loi des grands nombres. Mais cette conception butte évidemment sur les mêmes obstacles que ceux rencontrés par Jacob. Nécessairement, elle suppose un très grand nombre d'observations comme l'illustrent les figures 56, 57, 58: plus la période (mesuelle, trimestrielle, annuelle) est courte, mieux on distingue l'apparente normalité de la distribution:

> "[L]a "Loi de Fréquence de l'Erreur" [= loi normale][...] règne avec sérénité au milieu de la plus complète confusion. Plus est important [l'échantillon] plus parfaite est sa domination. C'est la surpême Loi de la Déraison. Quand un large échantillon d'éléments chaotiques sont mis ensemble [...] une régularité insoupçonnable et très belle se révèle." (Francis Galton, cité par [Sch64] p. 114)

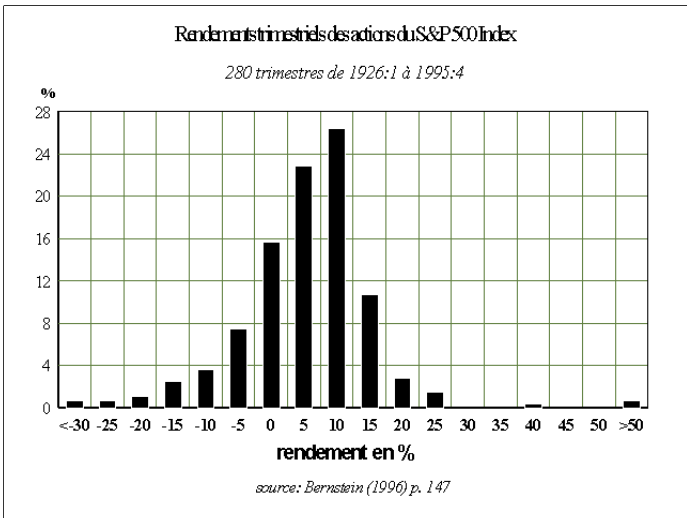
Une dernière école, l'école de la "propensation" (C.S. Peirce, Popper [Pop59]) a tenté de passer outre à cet obstacle en recourant à un argument métaphysique: la probabilité serait la disposition, la tendance de la Nature (!) à favoriser l'occurence de l'événement dans une expérience unique! Inutile d'ajouter que cette argumentation pour venir au secours de l'approche objective n'a pas rencontré un succès à la hauteur de sa poésie.

Cette vision cependant a le mérite d'attirer l'attention sur le problème posé par les expériences uniques. De nombreux événements aléatoires ne peuvent à la différence d'un lancement de dés être répliqués à volonté. Une course hippique, une coupe d'Europe de football constituent de tels exemples: au-

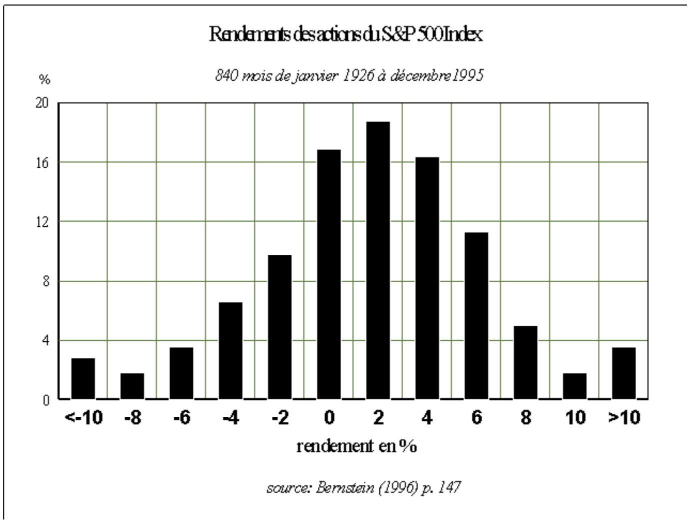
[27] Von Mises fut partiellement inspiré par les travaux des analyses de John Venn (1834-1923). Celui-ci était notamment l'auteur d'un ouvrage *Logic of Chance* paru en 1866 que Keynes décrivait comme "étonnament originale et qui influença considérablement le développement de la théorie de la statistique". Von Mises combina l'idée de Venn sur les valeurs limites avec la notion de séquence aléatoire d'un événement. L'approche de von Mises rencontra cependant peu de succès auprès des mathématiciens comme en témoigne l'éloge (sic!) posthume d'Alexander Ostrowski prononcé en 1965: "En raison de sa personnalité dynamique, ses bourdes occasionnelles étaient tolérées. On lui pardonna même sa théorie des probabilités." En fait, l'approche de von Mises fut victime de la révolution initiée par Andrei Kolmogorov (1903-1987). Cependant, celui-ci professait une vision appliquée proche de von Mises: "... la base des applications des résultats de la théorie mathématique des probabilités des 'phénomènes aléatoires' réels doit dépendre d'une conception fréquenciste des probabilités, la réalité incontournable qui a été argumentée avec fougue par von Mises." (déclaration de 1963)



Figure~56:



Figure~57:



Figure~58:

cune n'est semblable à une autre.[28] On voit donc se dessiner une distinction entre deux types d'événements: ceux relevant d'une approche objective, statistique et les autres. Ceci rejoint la célèbre distinction de Frank Knight (1885-1972) [Kni64] entre *risque* et *incertitude*:

> "L'Incertitude doit être dans un sens radicalement différent de la notion famillière de Risque, de laquelle elle n'a pourtant jamais été clairement distinguée. [...] Il apparaîtra qu'un incertitude *mesurable*, ou à proprement parler un "risque" [...] est totalement différent d'une incertitude *non mesurable*." ([Kni64] p. 205)

Dans l'interprétation de Knight (chapitre 7 p. 20 de [Kni64]), le "risque" est donc une situation où le décideur peut assigner une probabilité (= fréquence) objective à chaque état possible. Au contraire, l'"incertitude" est une situation où ceci est impossible. Keynes exprima des vues similaires dans son *Treatise on Probability* [Key21] et dans différents articles:

> "Par connaissance 'incertaine', [...] je ne veux simplement distinguer ce qui est connu avec certitude de ce qui est seulement possible. Le jeu de la roulette n'est pas soumis, dans ce sens, à l'incertain... [mais] les perpspectives de guerre en Europe sont incertaines, ou

[28] Même si les participants étaient nominalement les mêmes, réellement ils seraient différents par leurs âges, leurs conditions physiques, leur moral, le contexte de l'épreuve, etc...

Figure~59: Frank Knight

> encore le prix du cuivre ou les taux d'intérêt dans 20 ans le sont aussi... Sur ces matières il n'y aucune méthode scientifique pour déterminer des probabilités. Nous ne savons simplement pas." (J.M. Keynes, 1937, [Key37]) [29]

Cette vision knightienne de l'incertitude repose évidemment sur une vision réductrice des probabilités, réduites à n'être que le reflet de fréquences effectives. Aussi, de nombreux mathématiciens, philosophes et économistes ont soutenu que cette distinction étaient essentiellement d'ordre épistémologique et non ontologique: les probabilités peuvent aussi n'être que des "croyances" qui n'ont aucune relation avec la fréquence effective des événements.

Mais cette distinction entre "risque contre incertitude" souligne une limite de la théorie de l'utilité espérée à la von Neumann: comme les probabilités sont introduites en même temps que les éventualités et avant les préférences, ces probabilités ne sont pas de simples croyances subjectives. Peut-on construire une théorie où les probabilités, comme la fonction d'utilité élémentaire, ne soit pas un input du problème mais un output, un produit des préférences?

[29] Mais Keynes lui-même, plus loin dans le même texte, concluait: "Néanmoins, les nécessités de l'action et de la décision nous obligent cependant à négliger cette réalité embarassante et à nous comporter exactement comme si nous avions sous nos yeux un bon petit résumé benthamien des coûts et avantages de chaque issue, chacune associée à sa probabilité d'occurence exacte..." (J.M. Keynes, 1937, [Key37])

6.2 L'axiomatisation de Savage

Un ami de Keynes, Frank Ramsey (1903-1930), mathématicien-philosophe, ne partageait pas son opinion. Il présenta ses vues dans un essai *Truth and Probability* [Ram30] (rédigé en 1926, publié en 1931) où il prenait le contre-pied de Keynes en proposant une théorie subjective des probabilités, affirmant qu'une mesure de probabilité reposant sur les convictions, les croyances "déduit sa grandeur à la fois des désirs [utilité subjective] et des croyances [probabilités subjectives] [.]". Elle n'est donc pas déterminée par une connaissance objective, impersonnelle, commune, mais par un savoir propre à chaque individu. La probabilité est donc une donnée éminement subjective.

L'idée "subjectiviste" est intuitive mais peut-elle être mathématisée? Peut-on construire une mesure de ces croyances subjectives?

Ramsey, dans son papier, ne se contenta de démolir la théorie de Keynes, de proposer une autre conception des probabilités. Il suggéra également une méthode de dérivation des probabilités subjectives, méthode indépendamment proposée également par le grand statisticien italien Bruno de Finetti (1906-1985) [Fin31] [Fin37]. L'idée de Ramsey et de de Finetti est en fait d'inférer les croyances des individus de leurs comportements en utilisant une hypothèse courante en économie:

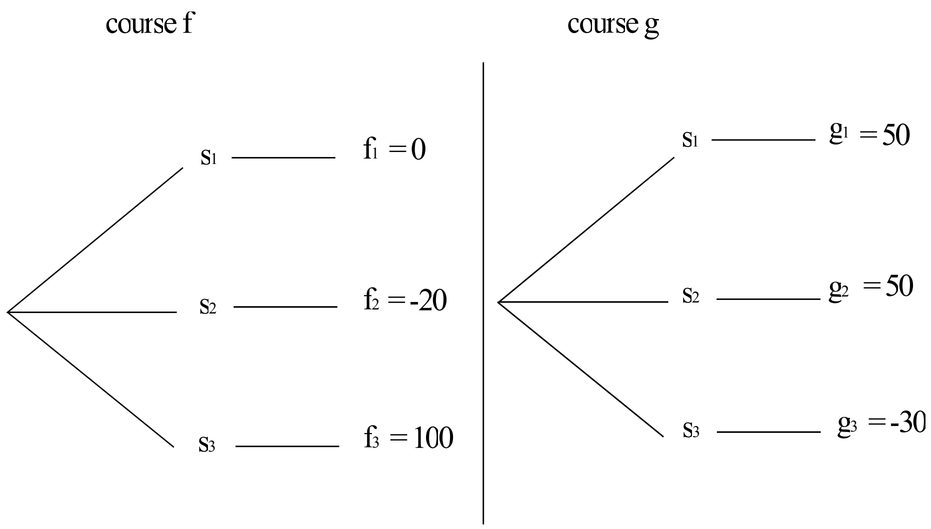
> "In order therefore to construct a theory of quantities of belief which shall be both general and more exact, I propose to take as a basis a general psychological theory, which is now universally discarded, but nevertheless comes, I think, fairly close to the truth in the sort of cases with which we are most concerned. I mean the theory that we act in the way we think most likely to realize the objects of our desires, so that a person's actions are completely determined by his desires and opinions. This theory cannot be made adequate to all the facts, but it seems to me a useful approximation to the truth particularly in the case of our self-conscious or professional life, and it is presupposed in a great deal of our thought. It is a simple theory and one which many psychologists would obviously like to preserve by introducing unconscious desires and unconscious opinions in order to bring it more into harmony with the facts. How far such fictions can achieve the required result I do not attempt to judge: I only claim for what follows approximate truth, or truth in relation to this artificial system of psychology, which like Newtonian mechanics can, I think, still be profitably used even though it is known to be false." ([Ram30] p. 173)

Figure~60: Frank Ramsey sur les bords du Red Pike dans le Lake District en 1925

Ainsi si une personne confrontée à deux résultats aléatoires x et y, le premier étant préféré au second. Supposons que l'agent soit désormais obligé de choisir entre deux loteries p et q définie sur les deux résultats. Si l'agent préfère la loterie p à la loterie q, nous pouvons en inférer que l'agent croit nécessairement que $p(x) > q(x)$. Ainsi, l'approche de Ramsey et de Finetti est une approche de probabilité révélée assez similaire à l'approche des préférences révélées.

Ce programme de dérivation trouva son prophète en Leonard J. Savage (1917-1971). Celui-ci dans son traité *Foundations of Statistics* (1954) exposa une dérivation complète des probabilités subjectives. Ce résultat a été regardé par les spécialistes (économistes et mathématiciens) comme une des constructions les remarquables de leurs disciplines respectives: "la plus brilliante des théories axiomatiques jamais développée" (Fishburn, [Fis70] p.191), "le couronnement de la théorie des choix" (Kreps, [Kre88] p.120).

Le cadre de la théorie de Savage est de supposer que l'incertain peut être représenté par un *ensemble d'états du monde* S , i.e. d'événements mutuellement incompatibles. Dans l'exemple de la course hippique, chaque état du monde est donc un classement possible des chevaux. A chaque état du monde possible correspond une conséquence, i.e. ici un prix, noté x. L'ensemble des prix possibles forme l'*ensemble des conséquences* X. Dans ce contexte, formellement une course hippique h est pour le parieur une fonction qui à chaque



Figure~61: Un exemple deux courses hippiques

état possible s associe un prix x:

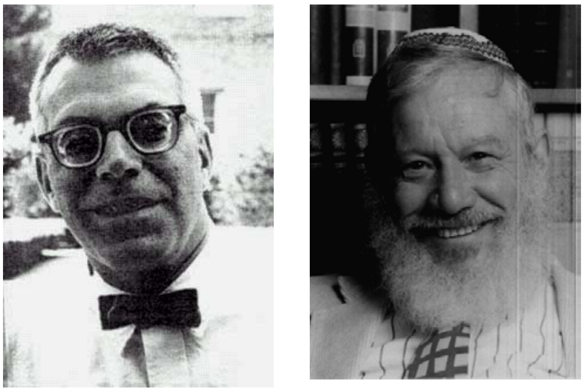
$$h : S \to X$$

Ainsi, sur la figure 61, une première course hippique f rapportera successivement 0, -20 et 100 alors que la course g, elle, donnerait 50, 50 et -30. L'*ensemble des actions possibles* est noté F:

$$F := \{f \mid f : S \to X\}$$

Evidemment une propriété fondamentale de ce cadre est que si l'on est capable de se représenter les issues possibles de chaque course, i.e. les états du monde, si l'on comprend parfaitement leurs conséquences, les prix, aucune probabilité n'apparaît: chacun est libre de se faire une croyance. Seul importe ici que chacun soit capable, quand on lui soumet deux paris sur une course possible, ou deux courses, par exemple f et g de la figure 61, de révéler s'il préfère strictement f à g, g à f, ou s'il est indifférent. Bref, comme toujours en économie, on suppose que les agents ont sur les décisions à prendre des préférences bien définies représentées par un pré-ordre complet.

Comme toute théorie, celle de Savage repose sur un corps d'hypothèses. Certaines sont techniquement assez similaires à celles de la théorie de l'utilité espérée à la von Neumann. Une hypothèse essentielle est l'*hypothèse de la chose sûre*:



Figure~62: Leonard J. Savage et Robert J. Aumann

Hypothèse 6 *Soit deux actions f, $g \in F$ telles que:*

$$f(s) = g(s) \ \forall s \in T$$

où T est un sous-ensemble non vide de S. Alors pour tout couple d'actions $(f', g') \in F \times F$ tel que:

$$f'(s) = g'(s) \ \forall s \in T$$

$$f'(s) = f(s) \ \forall s \in S \backslash T$$

$$g'(s) = g(s) \ \forall s \in S \backslash T$$

alors:

$$f \succ g \Leftrightarrow f' \succ g'$$

L'hypothèse de la chose sûre de Savage est dans ce contexte assez proche de l'hypothèse d'indépendance: elle assure en effet que seuls les différences joueront, i.e. les conséquences sur $S \backslash T$, que ces différences joueront quelles que soient les valeurs sur le complémentaire, i.e. que l'on prenne $f(s)$ ou $f'(s)$ sur T.

Avec cette hypothèse fondamentale (et d'autres), Savage a montré que pour tout pré-ordre $\succeq \subset F \times F$ il existe une fonction $u : X \rightarrow \Re$ et une probabilité $\pi \in \Delta(S)$ définie sur S telles que pour tout $(f, g) \in F \times F$:

$$f \succeq g \Leftrightarrow \sum_s \pi(s) u(f(s)) \geq \sum_s \pi(s) u(g(s))$$

Autrement dit, même s'il n'existe pas de probabilités objectives caractérisant les états du monde, lorsque les préférences satisfont certaines propriétés, alors il est possible de représenter les préférences par un indice d'utilité U en distinguant les préférences sur les conséquences, représentées par la fonction u, des croyances sur les états du monde, résumées par la probabilité π:

$$U(f) = \sum_s \pi(s)\, u(f(s))$$

La démonstration de Savage est une construction mathématique extrêmement compliquée et sollicite lourdement la théorie de la mesure ainsi que l'analyse fonctionnelle. Heureusement, près de 10 ans après la parution du *Foundation* de Savage, F.J. Anscombe et R.J. Aumann [AA63] démontrèrent un résultat proche en recourant à un cadre mathématiquement plus simple. Aussi, plutôt que de nous pencher sur la démonstration de Savage nous allons suivre celle de Anscombe & Aumann.

6.3 L'approche à la Anscombe & Aumann

L'axiomatisation de Savage est extrêmement lourde. Une présentation plus simple de l'utilité espérée subjective est celle présentée par Anscombe & Aumann [AA63]. L'astuce sur laquelle elle repose est de réintroduire dans le cadre les loteries que l'on avait chassé des préférences. En effet, à la différence du cadre de Savage, une action f n'est pas une fonction qui va de l'ensemble des états S dans l'ensemble des résultats X, mais:

$$f : X \to \Delta(X)$$

où $\Delta(X)$ est l'ensemble des probabilités (simples) définies sur X. Une conséquence n'est donc plus un x particulier, mais une distribution $p \in \Delta(X)$. L'ensemble des conséquences $\Delta(X)$ est lui-même un ensemble de loteries.

Les éléments du cadre d'Anscombe & Aumann sont donc les suivants:

- S l'ensemble des états;
- $\Delta(X)$ l'ensemble des conséquences;
- $f : S \to \Delta(X)$ une action;
- $F = \{f : S \to \Delta(X)\}$ l'ensemble des actions possibles;
- $\succeq \subset F \times F$ les préférences strictes sur les actions.

Une relation de préférence $\succeq$ est une relation binaire sur les actions F qui vérifie les axiomes suivants:

Axiome 4 $\succeq$ *est complète: ou* $f \succeq g$ *ou* $g \succeq f$ *pour tout* f , $g \in F$.

Axiome 5 $\succeq$ *est transitive:*

$$[f \succeq g, g \succeq h] \Rightarrow f \succeq h \forall f, g, h \in F$$

.

Axiome 6 *Soient trois loteries* f, g, $h \in F$ *vérifiant:*

$$f \succ g \succ h$$

alors il existe deux réels α, $\beta \in]0,1[$ *pour lesquels:*

$$\alpha f + (1-\alpha) h \succ g \succ \beta f + (1-\beta) h$$

Axiome 7 *Soient deux loteries* f, $g \in F$ *vérifiant:*

$$f \succ g$$

alors pour toute loterie $h \in F$, *pour tout réel* $\alpha \in]0,1[$ *on a:*

$$\alpha f + (1-\alpha) h \succ \alpha g + (1-\alpha) h$$

Les axiomes 4 - 7 sont évidemment similaires aux axiomes à la vNM. L'ensemble des actions est également comme dans ce cadre un ensemble convexe:

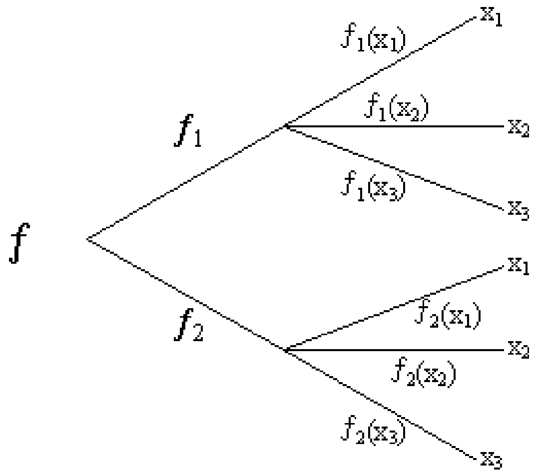
$$\forall f, g \in F, \forall a \in [0,1], af + (1-a)g \in F$$

où:

$$[af + (1-a)g](s) = af(s) + (1-a)g(s)$$

Intuitivement, l'interprétation du cadre à la Anscombe & Aumann est dans le contexte des courses hippiques la suivante: l'ensemble des états S est donc l'ensemble des vainqueurs possibles; $f: S \to \Delta(X)$ est la course hippique à laquelle peut participer notre parieur; mais, dans chaque état possible s, on gagne non un prix (élémentaire) mais un billet de loterie $f(s)$.[30] Sur la figure 63 est représentée une telle course hippique dans le cas où $S = \{1,2\}$ et $X = \{x_1, x_2, x_3\}$. Selon le vainqueur que couronnera Dame Nature, on recevra soit le billet f_1 soit le billet f_2. Chaque billet f_s donne le droit de recevoir le prix x_i avec une certaine probabilité $f_s(x_i)$.

[30] Dans les cas auquel on confrontera les décideurs, les loteries pourront *de fait* être des loteries dégénérées, donnant un prix avec certitude. Mais, *a priori*, pour caractériser les préférences, on élargit l'espace des conséquences.



Figure~63: Un exemple de loteries dans Anscombe & Aumann

Compte tenu des hypothèses, le théorème 3 sur l'utilité espérée nous assure que pour tout $s \in S$, il existe une fonction d'utilité espérée $U_s : \Delta(X) \rightarrow \Re$, avec:

$$U_s(f_s) = \sum_{x \in X} f_s(x_i)\, u_s(x_i)$$

où $u_s : X \rightarrow \Re$ est l'utilité élémentaire à la vNM dans l'état s. Ainsi, sur la figure 63, si l'état qui se réalise est $s = 1$, alors l'utilité espérée de f_1 est:

$$U_1(f_1) = f_1(x_1)u_1(x_1) + f_1(x_2)u_1(x_2) + f_1(x_3)u_1(x_3)$$

et si c'est l'état $s = 2$, l'utilité espérée de la loterie f_2 est:

$$U_2(f_2) = f_2(x_1)u_2(x_1) + f_2(x_2)u_2(x_2) + f_2(x_3)u_2(x_3)$$

Mais ces représentations ne portent que sur les conséquences et non sur les actions. Les préférences $\succeq$ sur les actions f, g, ..., peuvent-elles aussi être représentées sous la forme de fonction d'utilité voire d'utilité espérée?

Le premier résultat de l'approche d'Anscombe & Aumann répond par l'affirmative en montrant que les préférences ont pour représentation possible la forme suivante:

$$U(f) = \sum_{s \in S} U_s(f_s)$$

ou encore:

$$U(f) = \sum_{s \in S} \sum_{x \in X} f_s(x_i) u_s(x_i)$$

l'utilité u_s étant être déduite de U_s connaissant f_s.

Théorème 4 *Soit $S = [s_1, .., s_n]$, $\Delta(X)$ l'ensemble des probabilités simples définies sur X. Soit $\succeq$ une relation de préférence définie sur $F = \{f : f : S \to \Delta(X)\}$. Alors si $\succeq$ vérifie les axiomes 4, 5, 6, 7 , il existe un ensemble de fonctions $\{u_s : X \to \Re\}_{s \in S}$ tel que pour tout $f, g \in F$:*

$$f \succeq g \Leftrightarrow \sum_{s \in S} \sum_{x \in X} f_s(x_i) u_s(x) \geq \sum_{s \in S} \sum_{x \in X} g_s(x) u_s(x_i)$$

De plus, si $\{v_s : X \to R\}_{s \in S}$ est un autre ensemble de fonctions représentant ces préférences, alors il existe $b \geq 0$ et a_s tels que $v_s = bu_s + a_s$.

L'application des axiomes de VNM dans le cadre d'Anscombe & Aumann suffit donc pour obtenir un résultat de représentation très fort puisque les préférences obtenues sont déjà séparables. Cependant, au prix d'un renforcement substantiel (?) des préférences, on peut obtenir un résultat encore plus admirable. L'objectif suivant est en effet d'obtenir que les fonctions $u_s : X \to \Re$ soient les mêmes pour tous les états s.

Avant d'introduire l'hypothèse supplémentaire, posons tout d'abord une définition:

Définition 3 *Un état $s \in S$ est* négligeable *pour le pré-ordre $\succeq$ si:*

$$(f_1, .., f_{s-1}, p, f_{s+1}, .., f_n) \sim (f_1, .., f_{s-1}, q, f_{s+1}, .., f_n)$$

$\forall\ p, q \in \Delta(X)$, $\forall f \in F$.

Si tous les états étaient négligeables, le problème serait trivial et sans intérêt car toutes les actions seraient équivalentes. En effet, si tous les états étaient négligeables, alors pour tout $f = (f_1, f_2.., f_n)$:

$$(f_1, f_2.., f_n) \sim (f_{1'}\ , f_2, .., f_n) \sim (f_{1'}\ , f_{2'}, .., f_n) \sim .... \sim (f_{1'}, f_{2'}, .., f_{n'})$$

Mais, $(f_{1'}, f_{2'}, .., f_{n'})$ peut être n'importe quelle action $g \in F$. Par conséquent, s'il n'existe que des états néglieables alors nécessairement $f \sim g$ pour tout $g \in F$.

Pour éviter une telle trivialité, on introduit donc l'hypothèse suivante:

Hypothèse 7 *Il existe f , $g \in F$ pour lesquelles $f \succ g$.*

L'axiome fondamentale pour obtenir la réprésentation souhaitée est en fait le suivant:

Axiome 8 *Soit $s \in S$ un état non négligeable, $p, q \in \Delta(X)$. Si on a:*

$$(f_1, ..., f_{s-1}, p, f_{s+1}, .., f_n) \succeq (f_1, ..., f_{s-1}, q, f_{s+1}, .., f_n)$$

alors pour tout autre état non négligeable $t \in S$ on doit également observer:

$$(f_1, ..., f_{t-1}, p, f_{t+1}, .., f_n) \succeq (f_1, ..., f_{t-1}, q, f_{t+1}, .., f_n)$$

Dans le contexte des courses hippiques, si vous préférez gagner le billet de loterie p au billet de loterie q que le vainqueur soit votre cheval préféré ou non, que la course est été superbe ou non. Comme l'axiome d'indépendance de vNM, cet axiome conduit donc à minorer le contexte des loteries, à supposer une condition de séparabilité puisque que l'ordre de tout couple de loterie p et q sera indépendant des états. Aussi parlerons-nous d'axiome de non-contingence traduction maladroite de "state-independence". Avec cet axiome fort, le résultat remarquable obtenu est le suivant:

Théorème 5 *Soit $S = [s_1, .., s_n]$, $\Delta(X)$ l'ensemble des probabilités simples définies sur X. Soit $\succeq$ une relation de préférence définie sur $F = \{f : f : S \to \Delta(X)\}$. Alors si $\succeq$ vérifie les axiomes 4, 5, 6, 7, 8 et l'hypothèse 7, il existe une probabilité π définie sur S et une fonction $u : X \to \Re$ telles que pour tout $f, g \in F$:*

$$f \succeq g \Leftrightarrow \sum_{s \in S} \pi(s) \sum_{x \in X} f_s(x) u(x) \geq \sum_{s \in S} \pi(s) \sum_{x \in X} g_s(x) u(x)$$

S'il existe un autre couple (p', v) représentant les préférences alors il existe $b > 0$ et a pour lesquels on a: $v = bu + a$ et $p = p'$.

7 Limites et extensions de l'utilité espérée

Comme toute modélisation, la théorie de l'utilité espérée est nécessairement réductrice, ne peut prétendre qu'être une caricature (peut-être utile) de la réalité. Comme beaucoup de théories, elle est falsifiable. Aussi n'est-il pas étonnant que très tôt, elle fut soumise à un feu roulant de critiques à la fois théoriques et empiriques. Parmi ces dernières l'une des plus célèbres est celle de Allais [All53] laquelle repose sur une astucieuse expérience montrant certaines des limites empiriques de l'axiome d'indépendance.

On considère pour cela des loteries dont les prix possibles sont:

$$x_1 = 0\$,\ x_2 = 100\$,\ x_3 = 500\$$$

On offre alors un premier couple de loteries (résumées par leurs probabilités) p_1 et p_2:

$$p_1 = (0, 1, 0),\ p_2 = (0.01, 0.89, 0.1)$$

La loterie p_1 a évidemment comme qualité d'être une loterie assurant un revenu certain de 100 même si son revenu espéré est légèrement plus faible:

$$\begin{aligned} p_1 &\rightarrow \textit{espérance} = 100\$ \\ p_2 &\rightarrow \textit{espérance} = 139\$ \end{aligned}$$

Souvent, l'argument de la sûreté est décisif et donc:

$$p_1 \succ p_2$$

Une seconde série de loteries (avec les mêmes prix) est alors présentée. Les deux loteries q_1 et q_2 sont les suivantes:

$$q_1 = (0.89, 0.11, 0)\,,\ q_2 = (0.90, 0, 0.10)$$

Ces deux loteries sont donc risquées mais la loterie q_2 pour une probabilité de ne rien obtenir à peine supérieure (0.90 contre 0.89) permet de substituer un prix de 500$ au prix de 100$. Aussi, là aussi, très souvent les personnes auxquelles on soumet ce choix préfère la seconde loterie:

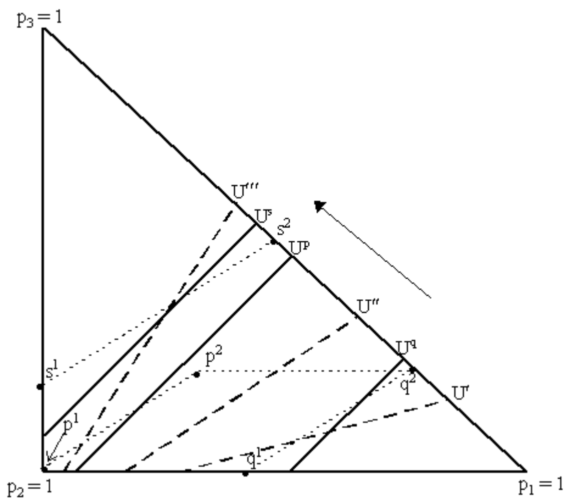
$$q_1 \prec q_2$$

Lorsqu'en 1951, au colloque international tenu à Paris réunissant les spécialistes de la recherche en théorie de la décision, Maurice Allais testa son expérience, certains partisans de l'utilité espérée, dont Leonard J. Savage, sélectionnèrent en effet:

$$p_1 \succ p_2,\ q_1 \prec q_2$$

Or, *ce choix est incompatible avec l'axiome d'indépendance!* Ceci peut être démontré soit par l'arithmétique, soit par un simple raisonnement graphique.

Sur la figure 64, dans le triangle de Marschak, sont représentées les loteries p_1, p_2, q_1, q_2. Graphiquement, une propriété fondamentale des loteries



Figure~64: Le paradoxe de Allais

concoctées par Maurice Allais devient évidente. Les loteries q_1 et q_2 ne sont des translations des deux premières:

$$\begin{aligned} q_1 &= p_1 + (0.89, -0.89, 0) \\ q_2 &= p_2 + (0.89, -0.89, 0) \end{aligned}$$

La translation, par construction, assure que les segments $[p_1, p_2]$ et $[q_1, q_2]$ sont parallèles . Or les courbes d'indifférences sont également parallèles si l'axiome d'indépendance est vérifié. Par conséquent, comme le montre la figure, si p_1 est séparé de p_2 par une courbe d'indifférence (U_p) alors nécessairement q_1 l'est également de q_2 par une autre (U_q). Les seuls choix (avec préférence stricte) compatibles avec l'axiome d'indépendance sont donc:

$$p_1 \succ p_2 \; et \; q_1 \succ q_2$$

et:

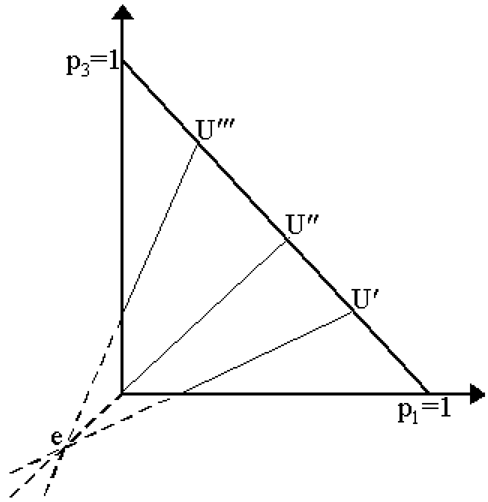
$$p_1 \prec p_2 \; et \; q_1 \prec q_2$$

Face au paradoxe de Allais, les partisans de l'utilité espérée (et notamment Savage et Marschak) ont adopté plusieurs arguments pour "sauver" l'utilité espérée. Un argument avancé essentiellement par Marschak [Mar51] consiste à la présenter comme un modèle normatif. Le paradoxe de Allais est donc alors tout simplement hors sujet. Mais l'importance de l'utilité espérée est également considérablement réduite.

Figure~65: Maurice Allais et une de ses ”victimes”, Leonard J. Savage

Un argument, avancé notamment par Savage, est qu'en décision, comme en calcul, les erreurs sont possibles lorsque le contexte est compliqué ou imparfaitement compris. Faire une erreur dans le calcul mental de $\frac{1160}{464}$ ne remet pas en cause l'arithmétique. Faire une erreur dans $\frac{1160}{464}$ ne conduit pas forcément à persévérer dans erreur lorsque l'on comprend enfin que $\frac{1160}{464} = \frac{5}{2} \times \frac{232}{232}$. Aussi, d'après Savage, ce que remet en cause l'expérience de Allais n'est pas tant l'axiome d'indépendance que l'expérience elle-même. Une expérience où le mode de construction des loteries serait une information commune, publique ne donnerait pas les mêmes résultats. De fait, dans les expériences à la Allais, la révélation de cette information conduit une part conséquente des personnes annonçant d'abord les choix “incohérents” à les corriger. Cependant, en général, au moins le quart des personnes conservent leurs choix “incohérents” initiaux.

Empiriquement, le paradoxe de Allais n'est pas un artefact mais est solidemment établi. Les préférences des agents semblent notamment violer l'axiome d'indépendance de manière systématique. En particulier, les courbes d'indifférence (révélées par les expériences) semblent se déployer en éventail (figure 64). Cette forme révèle une forte préférence pour la sécurité, un effet de prix de consolation selon Kahneman & Tversky [KT89]. Dans le choix $p_1$contre p_2, l'issue commune est avec une probabilité de 0.89 de gagner 100, les deux loteries divergeant sur la probabilité résiduelle de 0.11; dans le choix $q_1$contre q_2 l'issue commune est avec une probabilité de 0.89 de gagner 0, les deux loteries divergeant sur la probabilité résiduelle de 0.11. L'axiome d'indépendance implique que ce qui est décisif sont les différences des loteries sur la probabilité résiduelle de 0.11: l'issue commune n'a aucune importance; peu importe qu'avec une probabilité de 0.89, on gagne 100 ou 0. Pour Kahneman & Tversky, cette issue commune joue le rôle d'un prix de consolation: si sa valeur est faible, il sera de peu d'importance et les agents respecteront



Figure~66: Le déploiement en éventail des courbes d'indifférence révélé par le paradoxe de Allais

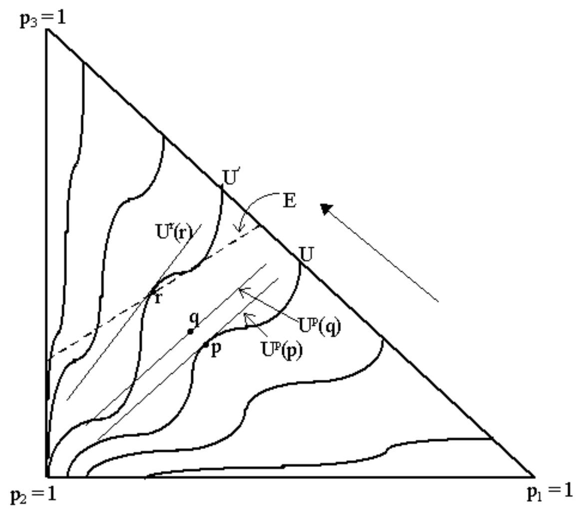
l'axiome d'indépendance en ne se souciant que des différences des loteries; si sa valeur est importante, les agents seront soucieux de ne pas faire de choix qu'ils pourraient regretter *ex post*, de ne pas lâcher la proie pour l'ombre. Or, dans le premier choix, le prix de consolation est de 100, et en sélectionnant la loterie p_1, on peut s'assurer de ce prix. Aussi, selon Kahneman & Tversky, le souci de s'emparer du prix de consolation conduit au choix de p_1.

Pour prendre en compte le paradoxe de Allais (et d'autres paradoxes), divers amendements de la théorie ont été proposés notamment à partir des années 80. L'une des plus intéressantes fut l'*utilité espérée non linéaire* de M. Machina [Mac82]. Dans cette approche, l'existence d'un pré-ordre complet et l'axiome archimédien est conservé mais non l'axiome d'indépendance. Les courbes d'indifférence ne sont plus en général ni parallèles, ni droites comme le montre la figure 67 où les prix sont supposés croissants: $x_1 < x_2 < x_3$.

L'idée centrale de la théorie de Machina est d'introduire une notion d'utilité espérée *locale*: en tout point, comme par exemple au point p de la figure, il doit être possible d'approximer la véritable courbe d'indifférence par la tangente en ce point. Formellement, cette approximation conduit à déterminer une fonction u^p spécifique à p approximant la véritable fonction d'utilité U:

$$U(q) \approx \sum_i q_i . u^p(x_i)$$

si q est suffisamment proche de p. Sous des hypothèses assez courantes (no-



Figure~67: Les courbes d'indifférence de la théorie de Machina

tamment de concavité des fonctions u^p), Machina a démontré que cette construction permettait de sauver la plupart des résultats essentiels obtenus avec l'utilité espérée tout en étant compatible avec les paradoxes à la Allais.

8 Annexe: démonstrations*

8.1 Théorème de Pratt

Théorème 6 *Soient u et v deux fonctions d'utilité élémentaires définies sur la richesse; les trois propriétés suivantes sont équivalentes :*
(i) $-\frac{u''}{u'} > -\frac{v''}{v'}$ *pour tout w;*
(ii) il existe une fonction croissance strictement concave G telle que:

$$u(w) = G(v(w))$$

(iii) si la richesse s'écrit $w + \widetilde{\epsilon}$, où w est une richesse certaine, $\widetilde{\epsilon}$ est un bruit blanc - $\mathbf{E}\widetilde{\epsilon} = 0$ *- alors:*

$$\rho_u(w) > \rho_v(w)$$

démonstration : $(i) \Rightarrow (ii)$: soit G la fonction implicite définie par $u(w) = G(v(w))$; les monotonicités des deux fonctions u et v nous assure que cette fonction est bien définie. Les dérivées de u s'écrivent:

$$\begin{aligned} u'(w) &= G'(v(w)).v'(w) \\ u''(w) &= G''(v(w)).v'^{\,2}(w) + G'(v(w))v''(w) \end{aligned}$$

Comme $u'(w)$ et $v'(w)$ sont strictement positifs, $G' > 0$. Le réarrangement de la deuxième équation donne :

$$\begin{aligned} v'(w)\frac{G''(v(w))}{G'(v(w))} &= \frac{u''(w)}{v'(w)G'(v(w))} - \frac{v''(w)}{v'(w)} \\ &= \frac{u''(w)}{u'(w)} - \frac{v''(w)}{v'(w)} < 0 \end{aligned}$$

Comme $v'(w) > 0$, $G'(w) > 0$, $G''(w) < 0$ et donc G est strictement concave.

(ii) $\Rightarrow$ *(iii)* : Pour établir cette implication, rappelons une relation fondamentale, l'inégalité de Jensen [Jen67]. Pour toute fonction strictement concave f la valeur de son espérance est supérieure à l'espérance de sa valeur:

$$f(\mathbf{E}\widetilde{w}) > \mathbf{E}f(\widetilde{w})$$

Par définition de la prime de risque, on a:

$$u(w - \rho_u(w)) = \mathbf{E}\left[u(w + \widetilde{\epsilon})\right]$$

Comme u est la composée de G et de v:

$$u(w - \rho_u(w)) = \mathbf{E}\left[G(v(w + \widetilde{\epsilon}))\right] \tag{14}$$

G est supposée une fonction strictement concave. Aussi, par l'inégalité de Jensen:

$$\mathbf{E}\left[G(v(w + \widetilde{\epsilon}))\right] < G\left(\mathbf{E}\left[v(w + \widetilde{\epsilon})\right]\right) \tag{15}$$

Comme la définition implicite de la prime de risque de v est:

$$\mathbf{E}\left[v(w + \widetilde{\epsilon})\right] = V(w - \rho_v(w)$$

on a:

$$G\left(\mathbf{E}\left[v(w + \widetilde{\epsilon})\right]\right) = G\left(V(w - \rho_v(w)\right) = u(w - \rho_v(w) \tag{16}$$

En combinant les relations (14), (15), (16), on obtient le résultat:

$$u(w - \rho_u(w)) < u(w - \rho_v(w)) \Rightarrow \rho_u(w) > \rho_v(w)\ \forall w$$

(iii) $\Rightarrow$ *(i)* : On se d'abord à une démonstration locale. Pour cela, on considère des bruits blancs obtenues par homothéties, i.e $\widetilde{\epsilon}(t) = t.\widetilde{\epsilon}$, et on paramétrise donc les primes d'assurance par le paramètre d'homothétie, $\rho_.(t, .)$. Un développement limité d'ordre 2 autour de 0 donne :

$$\rho(t, w) = \rho(0, w) + \rho'(0, w).t + \frac{t^2}{2}.\rho''(0, w)$$

La définition de ρ:

$$u(w-\rho(t,w)) = \mathbf{E}\left[u(w+t\widetilde{\epsilon})\right]$$

et la différenciation de cette équation donne successivement :

$$\begin{aligned} -u'(w-\rho(t,w))\rho'(t,w) &= \mathbf{E}[u'(w+t\widetilde{\epsilon})\widetilde{\epsilon}] \\ u''(w-\rho(t,w))\rho'(t,w)^2 - u'(w-\rho(t,w))\rho^{''}(t,w) &= \mathbf{E}[u''(w+t\widetilde{\epsilon})\widetilde{\epsilon}^2] \end{aligned}$$

De la première équation, on déduit: $\rho(0,w)=0$. En conséquence, la seconde expression donne :

$$\rho''(0,w) = -\frac{\mathbf{E}[u''(w)\widetilde{\epsilon}^2]}{u'(w)} = -\frac{u''(w)}{u'(w)}\mathbf{E}[\widetilde{\epsilon}^2] = r(w).\sigma_\epsilon^2$$

puisque la variance d'un bruit blanc σ_ϵ^2 est $\mathbf{E}[\widetilde{\epsilon}^2]$. Le développement limité donne donc:

$$\rho(0,w) = \left[-\frac{u''(w)}{u'(w)}\right]\frac{\sigma^2}{2}\times t^2 \qquad (17)$$

La démonstration complète est la suivante. La démonstration rigoureuse consiste à démontrer que "non (i)" $\Rightarrow$ "non (iii)". si "non (i)", alors il existe au moins un $w'>0$ tel que:

$$-u^{''}(w')/u'(w') < v''(w')/v'(w')$$

Par continuité, ceci est vrai dans un voisinage $N_\varepsilon(w')$ de w'. On prend alors une variable aléatoire z dont les seules valeurs positives sont des éléments de $N_\varepsilon(x')$.

Comme $-u^{''}(w)/u'(w) < v''(w)/v'(w)$ pour tout $w \in N_\varepsilon(w')$, la fonction G est cette fois strictement convexe (cf $(i) \Rightarrow (ii)$). Mais, G convexe, $G''>0$, implique nécessairement (cf $(ii) \Rightarrow (iii)$): $\rho^v(z) \geq \rho^u(z)$, i.e. "non (iii)". Thus, "non (i)" $\Rightarrow$ "non (iii)" ou, de la manière équivalente, (iii) $\Rightarrow$ (i). ■

8.2 Théorème d'utilité espérée

Théorème 7 *Sous les axiomes 3 et 2, les hypothèses 1, 2, 3, 5 et 4, pour tout pré-ordre complet* $\succeq$ *défini sur un ensemble £ il existe une fonction d'utilité U représentant ces préférences et satisfaisant la propriété d'espérance d'utilité:*

$$U\left(\left[(x_i)_{i=1,\dots}, (p_x(x_i))_{i=1,\dots}\right]\right) = \sum_i p_x(x_i).u(x_i)$$

où u est la fonction d'utilité élémentaire de U.

démonstration : En appliquant la démarche de Wold, on définit la fonction u de la manière suivante : $u(w) = 0$, $u(b) = 1$. Par le lemme 2, pour tout prix x il existe un unique réel $u(x) \in [0, 1]$ pour lequel:

$$x \sim u(x) * b \oplus (1 - u(x)) * w$$

Considérons une loterie $a_x = [(x_1, x_2), (p_x(x_1), p_x(x_2))]$. Naturellement:

$$x_1 \sim u(x_1) * b \oplus (1 - u(x_1)) * w$$

$$x_2 \sim u(x_2) * b \oplus (1 - u(x_2)) * w$$

Comme toute combinaison symétrique avec une troisième loterie r est équivalente, i.e. $\forall \alpha \in]0, 1[$:

$$\alpha * x_1 \oplus (1 - \alpha) * r \sim \alpha * (u(x_1) * b \oplus (1 - u(x_1))) \oplus (1 - \alpha) * r$$

Mais ceci vaut alors pour $\alpha = p_x(x_1)$, $r = x_2$ et donc:

$$a_x \sim p_x(x_1) * (u(x_1) * b \oplus (1 - u(x_1))) \oplus p_x(x_2) * (u(x_2) * b \oplus (1 - u(x_2)))$$

Le terme de droite est une loterie composée donnant droit soit à b soit à w. Par l'hypothèse 3 de réduction des loteries composées, en regroupant les termes facteurs de b et de w, la seconde loterie se réécrit:

$$(p_x(x_1) u(x_1) + p_x(x_2) u(x_2)) * b \oplus (1 - (p_x(x_1) u(x_1) + p_x(x_2) u(x_2)) * w$$

Par conséquent, l'agent est indifférent entre la loterie a_x et le fait de recevoir avec une probabilité $p_x(x_1) u(x_1) + p_x(x_2) u(x_2)$ la loterie b, avec la probabilité compémentaire la loterie w. L'utilité de a_x est donc bien donnée par l'utilité espérée de ses prix:

$$U(a_x) = p_x(x_1) u(x_1) + p_x(x_2) u(x_2)$$

Si la loterie comportait plus de deux prix, disons n, on obtiendrait le même résultat en décomposant la loterie initiale en une succession de n loteries (composées) emboitées les unes dans les autres:

$$a_x(1) = [(x_1, a_x(2)), (p_x(x_1), 1 - p_x(x_1))]$$

avec:

$$a_x(2) = \left[(x_2, a_x(3)), \left(\frac{p_x(x_2)}{1 - p_x(x_1)}, \frac{1 - p_x(x_1) - p_x(x_2)}{1 - p_x(x_1)}\right)\right]$$

$$a_x(n) = [x_n, 1]$$

A la première étape on montrerait comme plus haut que:

$$a_x \sim p_x(x_1) * (u(x_1) * b \oplus (1 - u(x_1))) \oplus (1 - p_x(x_1)) * a_x(2)$$

Puis on expliciterait $a_x(2)$, regouperait les termes en b et en w pour obtenir:

$$\begin{aligned} a_x \sim & (p_x(x_1) u(x_1) + p_x(x_2) u(x_2)) * b \oplus (1 - (p_x(x_1) u(x_1) + p_x(x_2) u(x_2)) w \\ & \oplus (1 - p_x(x_2)) * a_x(3) \end{aligned}$$

Itérant sur les loteries $a_x(3)$, $a_x(4)$, ..., on regrouperait successivement les termes pour obtenir finalement:

$$\begin{aligned} a_x \sim & \left(\sum_{i=1}^{n-1} p_x(x_i) u(x_i)\right) * b \oplus (1 - \sum_{i=1}^{n-1} p_x(x_i) u(x_i)) w \\ & \oplus \left(1 - \sum_{i=1}^{n-1} p_x(x_i)\right) * a_x(n) \end{aligned}$$

et donc avoir:

$$\begin{aligned} a_x \sim & \left(\sum_{i=1}^{n-1} p_x(x_i) u(x_i)\right) * b \oplus (1 - \sum_{i=1}^{n-1} p_x(x_i) u(x_i)) w \\ & \oplus p_x(x_n) [u(x_n) b + (1 - u(x_n) w)] \end{aligned}$$

d'où finalement après regroupement des termes:

$$a_x \sim \left(\sum_{i=1}^{n} p_x(x_i) u(x_i)\right) * b \oplus (1 - \sum_{i=1}^{n} p_x(x_i) u(x_i)) w$$

■

Références

[AA63] F.J. Anscombe and R.J. Aumann, (1963). A definition of subjective probability. *Annals of Mathematical Statistics*, 34: 199–205, 1963.

[All53] M. Allais, (1953). Le comportement de l'homme rationnel devant le risque, critique des postulats et axiomes de l'ecole américaine. *Econmetrica*, 21: 503–46, 1953.

[Bau51] W.J. Baumol, (1951). The neumann-morgenstern utility index: an ordinalist view. *Journal of Political Economy*, 59(1): 61–66, 1951.

[Bau58] W.J. Baumol, (1958). The cardinality utility which is ordinal. *Economic Journal*, 68: 665–72, 1958.

[Ber68] D. Bernoulli, (1738). Exposition of a new theory of risk evaluation. In W.J. Baumol & S.M. Goldfeld, editor, *Precursors in Mathematical Economics : an Anthology*, chapter 2, pages 15–26. The London School of Economics and Political Science, University of London, 1968. Réédition de la traduction de L. Sommers parue dans Econometrica.

[Ber95] P.L. Bernstein, (1995). *Des idées capitales*. Finance. PUF, Paris, 1995. Traduction de "Capital ideas - the improbable origins of modern Wall Street" (1992).

[Ber96] P.L. Bernstein, (1996). *Against the gods - The remarkable story of risk*. John Wiley & Sons, Inc., New York, 1996.

[Car30] J. Cardan, (1930). *Da Vita Propria Liber: The book of my life*. Dutton & Co, New York, 1930. Traduit du latin par Jean Stoner.

[Dav62] F.N. David, (1962). *Games, gods, and gambling*. Hafner Publishing Compagny, New York, 1962.

[Deb59] G. Debreu, (1959). *Théorie de la valeur*. Dunod, Paris, 1984, 2ème edition, 1959.

[Deb64] G. Debreu, (1964). Continuity properties of paretian utility. *International Economic Review*, 5: 285–93, 1964.

[Fin31] B. De Finetti, (1931). Sul significqto soggettivo dellq probabilita. *Fundamenta Mathematicae*, 17: 298–329, 1931.

[Fin37] B. De Finetti, (1937). La prévision: ses lois logiques, ses sources subjectives. *Annales de l'Institut Henri Poincaré*, 1937.

[Fis70] P.C. Fisburn, (1970). *Utility Theory for Decision Making*. John Wiley and Sons, New York, 1970.

[Hac75] I. Hacking, (1975). *The emergence of probability: a philosophical study of early studies about probability, induction and statistical inference*. Cambridge University Press, London, 1975.

[HM53] I. Herstein and J. Milnor, (1953). An axiomatic approach to measurable utility. *Econometrica*, 21: 291–97, 1953.

[Jen67] M.C. Jensen, (1967). Random walks: reality or myth - comment. *Financial Analysts Journal*, November-December 1967.

[Key21] J.M. Keynes, (1921). *A treatise on probability*. Macmillan, London, 1921.

[Key37] J.-M. Keynes, (1937). The general theory. *Quarterly Journal of Economics*, 51: 209–33, February 1937. Reproduit "The Collected Writing of John Maynard Keynes", vol. 14.

[Kni64] F.H. Knight, (1921). *Risk, uncertainty and profit*. Century Press, New York, 1964.

[Kre88] D.M. Kreps, (1988). *Notes on the theory of choice*. Westview Press, Boulder, Co., 1988.

[Kre90] D.M. Kreps, (1990). *A course in microeconomic theory*. Harvester Wheatsheaf, New York, 1ère edition, 1990.

[KT89] D. Kahneman and A. Tversky, (1989). Prospect theory: an analysis of decision under risk. *Econometrica*, 47: 263–91, 1989.

[Lap95] P.S. Laplace, (1795). *Essai philosophique sur les probabilités*. Paris, 1795.

[Lap12] P.S. Laplace, (1812). *Théorie analytique des probabilités*. Paris, 1812. Disponible sur le site Gallica de la BNP.

[Mac82] M. Machina, (1982). "expected utility" analysis without the independence axiom. *Econometrica*, 50: 277–323, 1982.

[Mac87] M. Machina, (1987). Choice under uncertainty: problems solved and unsolved. *Journal of Economic Perspectives*, 1(1): 121–54, 1987.

[Mal52] E. Malinvaud, (1952). Note on von neumann-morgenstern's strong independence axiom. *Econometrica*, 20: 679, 1952.

[Mal96] B.G. Malkiel, (1996). *A random walk down wall Street*. W.W. Norton & Compagny, New York, 6 ème edition, 1996.

[Mar50] J. Marschak, (1950). Rational behavior, uncertain prospects and measurable behavior. *Econometrica*, 18: 111–41, 1950.

[Mar51] J. Marschak, (1951). Why "should" statisticians and businessmen maximize moral expectation? In *Proceedings of the Second Berkeley Symposium on Mathematical Statistics and Probability.* University of California Press, Berkeley, 1951.

[Mis54] R. Mises, (1928). *Probability, statistics and truth.* Macmillan, New York, 1954.

[Pas62] B. Pascal, (1670). *Pensées.* Points. Seuil, 1962.

[Pop59] K.P. Popper, (1959). The propensity interpretation of probability. *British Journal of the Philosophy of Science*, 10: 25–42, 1959.

[Pra64] J. Pratt, (1964). Risk aversion in the small and in the large. *Econometrica*, 32: 122–36, 1964.

[Ram30] F. Ramsey, (1926). Truth and probability. In *Foundations of Mathematics and Other Logical Essays.* Routledge and Kegan Paul Ltd, Londres, 1930.

[Sam50] P.A. Samuelson, (1950). Probability and attempts to measure utility. *Economic Review*, 1(3): 167–73, 1950.

[Sam52] P.A. Samuelson, (1953). Utility, preferences, and the independence axiom. *Econometrica*, 20: 670–78, 1952.

[Sch64] W.L. Schaaf, (1964). *Carl Friedrich Gauss.* Franklin Watts, New York, 1964.

[VNM44] J. Von Neumann and O. Morgenstern, (1944). *Theory of Games and Economic Behavior.* Princeton University Press, Princeton, 1944.

[Wol43] H. Wold, (1943). A synthesis of pure demand analysis, i, ii, iii. *Skandinavisk Aktuarietidskrift*, 26, 27: 85–118, 221–63, (1944) 69–120, 1943.

[Wol52] H. Wold, (1952). Ordinal preferences or cardinal utility. *econometrica*, 20: 661–64, 1952.

[Yaa69] M. Yaari, (1969). Some measures of risk aversion. *Journal of Economic Theory*, 1: 315–29, 1969.