Probabilité du conditionnel et probabilité conditionnelle

M. Cozic (DEC, ENS) & P. Égré (CNRS, IJN)

Séance 6, 2 avril 2007

• E.W. Adams, "The Logic of Conditionals", *Inquiry*, vol. 8, 1965, pp. 166-97

• E.W. Adams, *A Primer of Probability Logic*, CSLI Publications, Stanford, 1998

• J. Bennett, *A Philosophical Guide to Conditionals*, Oxford UP, Oxford, 2003, **chap. 4-9**

• D. Edgington, "Conditionals", The Stanford Encyclopedia of Philosophy (Fall 2001 Edition), Edward N. Zalta (ed.), URL = `http://plato.stanford.edu/archives/fall2001/entries/conditionals/`, (§ 3 - The Suppositional Theory)

1 Probabilités et probabilités conditionnelles

1.1 Probabilités

• Le formalisme des probabilités est d'ordinaire exprimé dans un *cadre ensembliste* : on définit les probabilités comme une certaine classe de mesures sur un espace booléen (pour une probabilité finiment additive) ou sur un espace mesurable (pour une probabilité σ-additive).

• En logique philosophique, on exprime en général les probabilités dans un *cadre logique*: les fonctions de probabilités n'assignent pas un poids à des sous-ensembles d'un espaces d'états mais à des *formules* d'un langage formel. Si le langage en question est par exemple celui de la logique propositionnel, $\mathcal{L}$, alors les fonctions de probabilités sont certaines fonctions P du type $P : \mathcal{L} \rightarrow \mathbb{R}$.

Définition 1
Soit une fonction $P : \mathcal{L} \rightarrow \mathbb{R}$: P est une ***fonction de probabilité*** *ssi elle satisfait aux axiomes suivants où $\phi, \psi \in \mathcal{L}$ et $\models_{LP}$ désigne la relation de conséquence logique de la logique propositionnelle :*

(K1) $0 \leq P(\phi) \leq 1$
(K2) Si $\models_{LP} \phi$, *alors* $P(\phi) = 1$
(K3) Si $\phi \models_{LP} \psi$, *alors* $P(\phi) \leq P(\psi)$
(K4) Si $\{\phi, \psi\} \models_{LP} \perp$, *alors* $P(\phi \vee \psi) = P(\phi) + P(\psi)$

Remarque : la définition est bien analogue aux définitions usuelles des probabilités sur un espace booléen. Soit $(S, \mathcal{E})$ un espace booléen (i.e. S est un espace d'états et $\mathcal{E}$ est une algèbre de Boole) : (i) P est une fonction à valeur dans $[0, 1]$ (ii) si E et E' sont disjoints, alors $P(E \cup E') = P(E) + P(E')$ et (iii) $P(S) = 1$.

Exemple 1
On jette un dé équiprobable à 6 faces. On dispose de six formules atomiques $\phi_1, ..., \phi_6$, *chacune signifiant : "Le dé tombe sur la face* i*". La fonction de probabilité naturelle a les propriétés suivantes :*

- $P(\phi_i) = 1/6$
- $P(\phi_i \vee \phi_j)_{j \neq i} = 1/3$
- $P(\bigvee_{i=1}^{6} \phi_i) = 1$

• Voici quelques résultats attendus qui découlent immédiatement de la définition précédente :

Proposition 1
Soient $\phi, \psi \in \mathcal{L}$ *;*

1. $P(\phi) \leq P(\phi \vee \psi)$
2. $P(\neg\phi) = 1 - P(\phi)$
3. *Si* $\models_{LP} \neg\phi$, *alors* $P(\phi) = 0$
4. *Si* $\phi \equiv \psi$, *alors* $P(\phi) = P(\psi)$
5. $P(\phi) + P(\psi) = P(\phi \vee \psi) + P(\phi \wedge \psi)$
6. $P(\phi) = P(\phi \wedge \psi) + P(\phi \wedge \neg\psi)$ *(théorème de l'addition)*

Preuve : 1. $\phi \models_{LP} (\phi \vee \psi)$ donc en vertu de (K3), $P(\phi) \leq P(\phi \vee \psi)$. 2. ϕ et $\neg\phi$ sont logiquement incompatibles donc $P(\phi \vee \neg\phi) = P(\phi) + P(\neg\phi)$. Or $\models_{LP} (\phi \vee \neg\phi)$ donc par (K2), $P(\phi \vee \neg\phi) = 1$. Donc $P(\neg\phi) = 1 - P(\phi)$, *etc.*

1.2 Probabilité conditionnelle

• Pour une fonction de probabilité P donnée, on note $P(\psi|\phi)$ la probabilité de ψ étant donné ϕ. On peut considérer que P est la probabilité **initiale** ou **a priori**: $P(\psi)$ est la probabilité initiale que l'on attache au fait que ψ. Supposons maintenant que ϕ soit le cas ; on peut concevoir $P(\psi|\phi)$ comme la nouvelle probabilité que l'on attache à ψ, probabilité qui tient du compte du fait que ϕ est le cas. On parle alors de probabilité **a posteriori** que ψ. Par conséquent, on peut voir les probabilités conditionnelles comme une représentation de la **révision des probabilités** qui fait passer de $P(.)$ à $P(.|\phi)$.

• Comment caractériser la probabilité conditionnelle $P(.|\phi)$ à l'aide de la probabilité initiale $P(.)$? La réponse classique est: par la **règle de Bayes**.

1. Considérons l'exemple du dé à six faces et supposons que l'on soit informé que le dé est tombé sur une face paire $((\phi_2 \vee \phi_4 \vee \phi_6))$. Question: quelle est la probabilité (*a posteriori*) pour que le dé soit tombé sur la face 2 (ϕ_2) ? Une réponse naturelle est la suivante: si le dé est tombé sur une face paire, alors il y a une probabilité de 1/3 pour qu'il soit tombé sur 2. En effet, il reste trois possibilités (2, 4 et 6), et chaque possibilité est (*a priori*) équiprobable. On peut décomposer le raisonnement comme suit: on prend la probabilité initiale de ϕ_2, et on la divise par la probabilité de $(\phi_2 \vee \phi_4 \vee \phi_6)$. Cela suggère le principe suivant : $P(\phi_2|(\phi_2 \vee \phi_4 \vee \phi_6)) = P(\phi_2)/P(\phi_2 \vee \phi_4 \vee \phi_6)$.

2. Ce n'est pas assez général. Considérons la même information, mais intéressons-nous désormais à la question suivante : quelle est la probabilité du fait que le dé soit tombé sur 2 ou 3 étant donné qu'il est tombé sur une face paire ? La réponse naturelle est de nouveau 1/3. On se rend compte que la formule précédente ne peut rendre compte de l'intuition, et à juste titre : il n'est pas possible que le dé soit tombé sur une face paire et qu'il tombe sur 3. Il faut donc éliminer la possibilité que le dé soit tombé sur 3. Il faut considérer la conjonction de $(\phi_2 \vee \phi_3)$ et de $(\phi_2 \vee \phi_4 \vee \phi_6)$. Cela suggère le principe suivant : $P(\phi_2 \vee \phi_3)|(\phi_2 \vee \phi_4 \vee \phi_6)) = P((\phi_2 \vee \phi_3) \wedge (\phi_2 \vee \phi_4 \vee \phi_6))/P(\phi_2 \vee \phi_4 \vee \phi_6)$.

3. La règle de Bayes est donnée par T. Bayes (1763) et reformulée d'un point de vue subjectiviste par Ramsey (1926) :

   > F.P. Ramsey, "Truth and Probability" (1926) trad.fr. "Vérité et probabilité"
   > 3ème loi fondamentale de la croyance probable :
   > (3) Degré de croyance en (p et q) = degré de croyance en p × degré de croyance en q étant donné p

   En d'autres termes,

$$\boxed{P(\psi|\phi) = P(\phi \wedge \psi)/P(\phi)}$$

Exemple 2
Reprenons l'exemple précédent. Supposons que l'on sache que le dé est tombé sur un nombre pair. Par conséquent, on sait qu'il est le cas que $(\phi_2 \vee \phi_4 \vee \phi_6)$*. On peut alors se demander*

- *quelle est la probabilité pour que le dé soit tombé sur la face 2 étant donné qu'il est tombé sur une face paire :*
  $P(\phi_2|(\phi_2 \vee \phi_4 \vee \phi_6))$
  $= P(\phi_2 \wedge (\phi_2 \vee \phi_4 \vee \phi_6))/P((\phi_2 \vee \phi_4 \vee \phi_6))$
  $= P(\phi_2)/P((\phi_2 \vee \phi_4 \vee \phi_6))$
  $= (1/6)/(1/2) = 1/3$
- *quelle est la probabilité pour que le dé soit tombé sur la face 2 ou la face 4 étant donné qu'il est tombé sur une face paire:*
  $P((\phi_2 \vee \phi_4)|(\phi_2 \vee \phi_4 \vee \phi_6)) = 2/3$
- *quelle est la probabilité pour que le dé soit tombé sur la face 2 ou la face 3 étant donné qu'il est tombé sur une face paire :*
  $P((\phi_2 \vee \phi_3)|(\phi_2 \vee \phi_4 \vee \phi_6)) = P(\phi_2|(\phi_2 \vee \phi_4 \vee \phi_6)) = 1/3$

• La règle de Bayes (ou "*ratio formula*") fait l'objet de nombreuses discussions conceptuelles : est-ce une définition de la probabilité conditionnelle? Est-ce une propriété de la probabilité conditionnelle? Certaines de ses caractéristiques sont également discutées et en particulier le fait qu'*elle n'est pas définie quand* $P(\phi) = 0$.

Pour "simplifier la théorie", Adams postule que lorsque $P(\phi) = 0$, pour tout formule ψ, $P(\psi|\phi) = 1$.

Proposition 2
Si l'on note $p(.|\psi)$ $p_\psi(.)$*, alors on démontre les propriétés suivantes :*

1. $0 \leq p_\psi(\phi) \leq 1$
2. *Si* $\psi \models_{LP} \phi$*, alors* $p_\psi(\phi) = 1$
3. *Si* $(\phi \wedge \psi) \models_{LP} \chi$*, alors* $p_\psi(\phi) \leq p_\psi(\chi)$
4. *Si* $p(\psi) > 0$ *et* $\{(\phi \wedge \psi), (\chi \wedge \psi)\} \models_{LP} \bot$*, alors* $p_\psi(\phi \vee \chi) = p_\psi(\phi) + p_\psi(\chi)$*.*

Remarque sur les conditionnels indicatifs vs contrefactuels: le fait que la probabilité conditionnelle soit non-définie en principe lorsque la probabilité de l'antécédent est nulle est considéré par certains comme l'indication d'un lien profond entre probabilité conditionnelle et la notion de conditionnel indicatif:

(1) Si Marie ne part pas demain, alors Pierre sera content.

(2) Si Napoléon n'avait pas perdu à Waterloo, la France serait un Empire.

L'une des propriétés du conditionnel indicatif, par opposition au conditionnel contre-factuel, c'est (généralement) qu'on considère l'antécédent du conditionnel comme possible, ce qui revient à dire qu'on lui associe une probabilité non-nulle (cf Bennett §23: "Indicative Conditionals are Zero-Intolerant"). Au contraire, pour un authentique contrefactuel, par définition, la probabilité de l'antécédent est nulle.

2 La thèse d'Adams

• Le **test de Ramsey**, derechef :

> F.P. Ramsey, 1929, "Law and Causality", trad.fr. J. Leroux "Les propositions générales et la causalité" dans *Logique, philosophie et probabilités*, Vrin, 2003, p. 246 :
>
> "If two people are arguing "If p will q ?" and are both in doubts as to p, they are adding p hypothetically to their stock of knowledge and arguing on that basis about q ;... they are fixing their degrees of belief in q given p"
>
> "Si deux personnes débattent de la question "Si p, est-ce que q ?" et qu'elles ont toutes les deux des doutes sur p, elles ajoutent p de façon hypothétique à leur stock respectif de connaissances et argumentent pour ou contre q sur cette base... *On peut dire que ces deux personnes sont en train de fixer leur degré de croyance en q, étant donné p*" [tr. modifiée]

• Dans un article de 1965, E. Adams propose pour la première fois l'une des thèses centrales de la théorie des conditionnels ; voici comment il l'exprime dans la dernière exposition :

> E.W. Adams, *A Primer of Probability Logic*, CSLI Publications, Stanford, 1998
>
> "we are hypothesizing that the *probabilities of conditional statements like "If Jane does take ethics, then she will take logic" are conditional probabilities.* This is the *probability conditional theory.*" (p. 114)

Cette thèse, nous l'appellerons la **thèse d'Adams** et nous l'exprimerons de la façon suivante :

$$\boxed{P(\text{“Si } \phi\text{, alors } \psi\text{”}) = P(\psi|\phi)}$$

• Que signifie exactement la thèse d'Adams ? Chez Ramsey, il s'agit de *fixation du degré de croyance*. Ce qui est présupposé est la conception *subjectiviste ou bayésienne* des probabilités selon laquelle les degrés de croyance d'un agent rationnel obéissent aux axiomes des probabilités. Cela signifie, par exemple, que si ϕ et ψ sont incompatibles, et si Pierre est un agent rationel, alors la somme du degré de croyance de Pierre en ϕ et de son degré de croyance en ψ doit être égale au degré de croyance en $(\phi \vee \psi)$. L'une des principales défenses de cette thèse est fournie par Ramsey dans “Vérité et probabilité” (1926). On la présente souvent en termes de **Dutch Book** (= marché de dupes).

> F.P. Ramsey, “Vérité et probabilité”, trad.fr. dans *Logique, philosophie et probabilité*, Vrin, 2003, p. 174
>
> “N'importe quel ensemble défini de degrés de croyance qui les violerait serait incohérent au sens où il violerait les lois de préférence entre les options.”

Indépendamment et peu de temps après, B. De Finetti (1937) a proposé un argument comparable basé sur le gain monétaire et non sur l'utilité. Voir D. Gillies, *Philosophical Theories of Probability*, Routledge, 2000, chap. 4

> B. De Finetti, “La prévision : ses lois logiques, ses sources subjectives”, *Annales de l'Institut H. Poincaré*, tome 7, N°1, 1937
>
> “...[le] calcul [des probabilités] apparaît comme l'ensemble des règles auxquelles l'évaluation subjective des probabilités de divers événements par un même individu doive être assujettie si l'on ne veut pas qu'il y ait entre elles de contradictions fondamentales.”

• Si l'on se place dans le cadre bayésien ou subjectiviste, alors la thèse d'Adams apparaît comme une thèse portant sur *l'épistémologie des conditionnels*. Elle affirme que le degré de croyance (rationnel) en “Si ϕ, alors ψ” est le degré de croyance en ϕ étant donné que ψ.

• *A priori*, la thèse d'Adams n'exclut pas que “Si ϕ, alors ψ” soit représenté par le conditionnel matériel, ie $(\phi \to \psi)$. Elle pose plutôt une contrainte épistémique sur une théorie des conditionnels. Est-il donc possible que $\to$ obéisse à la thèse d'Adams, soit $P(\phi \to \psi) = P(\psi|\phi)$?

Exemple 3
Reprenons l'exemple du jet de dé. Nous avons vu que $P(\phi_2|(\phi_2 \vee \phi_4 \vee \phi_6)) = 1/3$*. Considérons maintenant*
$P((\phi_2 \vee \phi_4 \vee \phi_6) \rightarrow \phi_2)$
$= P(\neg(\phi_2 \vee \phi_4 \vee \phi_6) \vee \phi_2)$
$= 2/3$

Il y a des cas particuliers où $P(\phi \rightarrow \psi) = P(\psi|\phi)$:

1. Si $P(\phi \wedge \neg\psi) = 0$, alors $P(\phi \rightarrow \psi) = P(\psi|\phi) = 1$
2. Si $P(\phi) = 1$, alors $P(\phi \rightarrow \psi) = P(\psi|\phi)$

Dans le cas général, comme le suggère l'exemple du jet de dé:

$$P(\phi \rightarrow \psi) \ \geq \ P(\psi|\phi)$$

$P(\phi \wedge \neg\psi)/P(\phi) \geq P(\phi \wedge \neg\psi)$
$P(\neg\psi|\phi) \geq P(\phi \wedge \neg\psi)$
$1 - P(\psi|\phi) \geq 1 - P(\phi \rightarrow \psi)$
$P(\phi \rightarrow \psi) \geq P(\psi|\phi)$

- Pourquoi accepter la thèse d'Adams ? Quelques raisons :

1. La thèse d'Adams "rend justice à la grammaire des énoncés de probabilité conditionnelle" (A. Hajek) : les probabilités de conditionnels et les probabilités conditionnelles semblent quasi-synonymes.

   B. van Fraassen, "Probabilities of Conditionals", dans Harper et Hooker (eds.), *Foundations of Probability Theory, Statistical Inference and Statistical Theories of Science*, vol.I, Reidel, 1976

   "The English statement of a conditional probability sounds exactly like that of the probability of a conditional. What is the probability that I throw a six if I throw an even number, if not the probability that: If I throw an even number, it will be a six ?"

2. On peut baser sur cette thèse une théorie de l'inférence probable qui ne retient pas les "mauvaises inférences" qui sont valides en logique classique.

"la plupart" et les probabilités conditionnelles: le lien entre la grammaire des conditionnels et la notion de probabilité conditionnelle est lié à ce que nous avons déjà vu en exercice sur la restriction des quantificateurs.

• Rappelons-nous qu'on ne peut pas interpréter: "la plupart des hommes, s'ils sont philosophes, sont sages", comme $[MOSTx : Hx][(Px \rightarrow Sx)]$: l'énoncé devrait être vrai dans une situation dans laquelle il y a 4 hommes, 3 non-philosophes, et 1 philosophe qui n'est pas sage.

En revanche, on se souvient que: $[MOSTx : Ax][Bx] \models MOSTx(Ax \rightarrow Bx)$. Le résultat est comparable à: $P(B|A) \leq P(A \rightarrow B)$. Il y a une analogie réelle ce point de vue entre:

(3) Si c'est un A, alors probablement c'est un B

(4) La plupart des x, s'ils sont A, sont B.

(5) Il y a 1 chance sur 2 pour que, si c'est A, alors ce soit un B

exercice: cf l'exercice, qui vous demande d'approfondir ce point.

3 Logique de la probabilité conditionnelle, I

3.1 Morphologie

• Si l'on accepte la thèse d'Adams, le conditionnel matériel ne peut refléter adéquatement le conditionnel. On va donc introduire un nouveau symbole pour le conditionnel : $\Rightarrow$. Par hypothèse, on aura le principe suivant :

$$P(\phi \Rightarrow \psi) = P(\psi|\phi) \tag{6}$$

• On ne va pas construire de langage formel avec $\Rightarrow$ comme on le fait avec les autres connecteurs (et comme on l'a fait avec les conditionnels de Stalnaker et de Lewis). La raison en est la suivante. Si $\Rightarrow$ était un connecteur comme les autres, on pourrait composer des formules à partir de conditionnels, ie on pourrait composer des formules à partir de sous-formules du types $(\phi \Rightarrow \psi)$. On pourrait par exemple former $\neg(\phi \Rightarrow \psi)$. Le problème, c'est que, du point de vue de l'interprétation probabiliste, on aurait alors quelque chose comme $P(\neg(\psi|\phi))$, et que l'on a pas de règles pour attribuer une probabilité à $\neg(\psi|\phi)$. La solution la plus simple au problème, que suit Adams (1998), consiste à n'autoriser l'occurrence de $\Rightarrow$ que comme connecteur principal.

Définition 2
Soit $\mathcal{F}(\mathcal{L})$ l'ensemble des formules d'un langage propositionnel $\mathcal{L}$; l'ensemble des formules du langage d'Adams pour le conditionnel, $\mathcal{F}(\mathcal{L})_{\Rightarrow}$ est
$\{\phi : \phi \in \mathcal{F}(\mathcal{L})\} \cup \{(\phi \Rightarrow \psi) : \phi, \psi \in \mathcal{F}(\mathcal{L})\}$

• *Remarque :* Adams appelle **factuelles** les formules de la logique propositionnelle $(\mathcal{F}(\mathcal{L}))$; et **conditionnelles** les formules qui sont formées par deux formules factuelles et $\Rightarrow$.

3.2 Validité probabiliste : premiers pas

> "..what many regard as the main advantage of the probability conditional theory [is] the way it resolves certain "paradoxes" of material conditionals.

• On a vu que la logique "expliquait" la notion intuitive d'argument valide par la relation de conséquence logique : un ensemble de prémisses a pour conséquence logique une conclusion ssi dès que toutes les prémisses sont vraies, la conclusion l'est également. Comme on le dit parfois, un argument valide satisfait une propriété de *préservation de la vérité*.

• Adams ne propose pas d'étendre la logique propositionnelle et d'étendre parallèlement la relation de conséquence logique (la façon dont on a procédé dans la théorie Stalnaker-Lewis); il propose une autre notion d'argument valide : une notion de **validité probabiliste**. Intuitivement (et en simplifiant), un argument sera qualifié de valide du point de vue probabiliste s'il n'est pas possible que la conclusion soit moins probable que les prémisses.

• L'une des ambitions de cette construction est de "résoudre" les paradoxes du conditionnel matériel et, plus généralement, de ne pas valider les inférences problématiques (vérité du conséquent, fausseté de l'antécédent, monotonie, transitivité, contraposition, etc.) On procède en deux étapes :

1. on substitue $\Rightarrow$ à $\rightarrow$ (on peut montrer que si l'on conserve $\rightarrow$, on ne pourra pas rejeter les inférences problématiques, voir ci-après)
2. on détermine si l'argument formulé en termes de $\Rightarrow$ est p-valide.

(1) **Paradoxe de la fausseté de l'antécédent :** $\neg\phi \models_{LP} (\phi \rightarrow \psi)$

Exemple 4
On tire deux cartes sans remplacement d'un paquet de 52 cartes ; à chaque tirage, les cartes qui peuvent être tirées ont toutes la même probabilité de l'être.
"La première carte sera un as" $\rightsquigarrow p$
"La seconde carte sera un as" $\rightsquigarrow q$
$P(\neg p) = 12/13 \simeq 0.923$
$P(p \Rightarrow q) = P(q|p) = 3/51 \simeq 0.059$
$P(p \Rightarrow \neg p) = P(\neg p|p) = 0$

L'exemple montre que l'on peut attacher une forte probabilité à $\neg\phi$ et une très faible probabilité à $(\phi \Rightarrow \psi)$. En un sens intuitif, la fausseté de l'antécédent n'est donc pas un argument valide du point de vue probabiliste.

(2) **Paradoxes de la vérité du conséquent:** $\psi \models_{LP} (\phi \rightarrow \psi)$

Exemple 5
"Il y aura cours de logique lundi prochain" $\rightsquigarrow q$
"Les enseignants seront malades lundi prochain" $\rightsquigarrow p$
$P(p) = 0.1$, $P(q) = 0.9$, $P(p \wedge q) = 0.01$, $P(\neg p \wedge \neg q) = 0.01$
$P(p \Rightarrow q) = P(q|p) = 0.1$
Il se peut donc parfaitement qu'il soit très probable qu'il y aura cours de logique lundi prochain mais qu'il soit très peu probable que si les enseignants sont malades lundi prochain, il y aura cours de logique ce même jour.

(3) **Contraposition :** $(\phi \rightarrow \psi) \models_{LP} (\neg\psi \rightarrow \neg\phi)$

Exemple 6
"Le soleil se lèvera demain" $\rightsquigarrow p$
"La température s'élèvera durant la journée $\rightsquigarrow q$
$P(p) = 0.99$, $P(q) = 0.901$, $P(p \wedge q) = 0.90$, $P(\neg p \wedge \neg q) = 0.009$
"Si le soleil se lève demain, alors la température s'élèvera durant la journée" $\rightsquigarrow P(p \Rightarrow q) = 0.909$
"Si la température ne s'élève pas durant la journée de demain, alors le soleil ne se lèvera pas demain" $\rightsquigarrow P(\neg q \Rightarrow \neg p) = 0.091$

(4) **Monotonie :** $(\phi \rightarrow \chi) \models_{LP} ((\phi \wedge \psi) \rightarrow \chi)$

Exemple 7
"Pierre mettra de l'essence dans son café" $\rightsquigarrow p$
"Pierre mettra du sucre dans son café " $\rightsquigarrow q$
"Pierre appréciera son café" $\rightsquigarrow r$
$P(p) = 0.09001$, $P(q) = 0.9909$, $P(r) = 0.90009$, $P(p \wedge q) = 0.09$, $P(q \wedge r) = 0.9$, $P(p \wedge r) = 0$ *et* $P(\neg p \wedge \neg q \wedge \neg r) = 0.009$

"Si Pierre met du sucre dans son café, il l'appréciera" $\rightsquigarrow P(q \Rightarrow r) = 0.9083$
"Si Pierre met du sucre et de l'essence dans son café, il l'appréciera" $\rightsquigarrow P((p \wedge q) \Rightarrow r) = 0$

(5) **Transitivité** : $\{(\phi \rightarrow \psi), (\psi \rightarrow \chi)\} \models_{LP} (\phi \rightarrow \chi)$

Exemple 8
"Le soleil se lèvera exactement 5 minutes plus tard demain" $\rightsquigarrow p$
"Le soleil ne se lèvera pas plus de 5 minutes plus tard demain" $\rightsquigarrow q$
"Le soleil se lèvera moins de 5 minutes plus tard demain" $\rightsquigarrow r$

On conserve les probabilités de l'exemple précédent. On vérifie que l'on a les probabilités conditionnelles suivantes :
$P(p \Rightarrow q) = 0.9999$, $P(q \Rightarrow r) = 0.9083$ *mais* $P(p \Rightarrow r) = 0$

En d'autres termes, il peut

- *être très probable* *que si le soleil se lève exactement 5 minutes plus tard demain, il ne se lèvera pas plus de 5 minutes plus tard*
- *être très probable* *que si le soleil ne se lève pas plus de 5 minutes plus tard demain, il se lèvera moins de 5 minutes plus tard*
- *être très peu probable* *que si le soleil se lève exactement 5 minutes plus tard demain, il se lèvera moins de 5 minutes plus tard*

• Du $\Rightarrow$ au $\rightarrow$: considérons l'inférence de $(\phi \Rightarrow \psi)$ à $(\phi \rightarrow \psi)$. Nous avons montré que, dans le cas général, $P(\phi \rightarrow \psi) \geq P(\psi|\phi)$; l'argument est par conséquent valide du point de vue probabiliste.