teoria bayesiana cattura un principio essenziale della metodologia popperiana: quello delle “ipotesi audaci” e dei “severi controlli” (vedi cap. 3).

Pertanto, attraverso il Teorema di Bayes, conoscendo i valori di $pr(H)$, di $pr(E \mid H)$ e di $pr(E)$, è possibile calcolare il valore di $pr(H \mid E)$, e, quindi, determinare se l’evidenza osservativa disponibile *E conferma*, *sconferma* o è *neutrale* rispetto da *H*: ove, nell’approccio bayesiano le nozioni epistemiche qualitative di «conferma», «sconferma» e «neutralità» sono definite (esplicate), in termini quantitativi rigorosi, come segue:

- *E conferma H* se e solo se $pr(H \mid E) > pr(H)$;
- *E sconferma H* se e solo se $pr(H \mid E) < pr(H)$;
- *E* è *neutrale* rispetto ad *H* se e solo se $pr(H \mid E) = pr(H)$.

Ne segue che la misura del *grado di conferma*, che una evidenza osservativa *E* fornisce a un’ipotesi *H*, è data da $pr(H \mid E) - pr(H)$.

4.2. Il criterio di confermabilità bayesiana

La nozione pragmatica (o epistemica) di *confermabilità* è stata spesso interpretata – come quella di verificabilità (vedi cap. 2) – sia come *criterio di significato* che come *criterio di demarcazione* (vedi, per es., Hempel, 1945 e Carnap, 1936-37 e 1950). Ma come criterio di significato essa si trova esposta alle stesse obiezioni che, nella sezione 2.1, abbiamo rivolto al criterio verificabilista del significato. Pertanto, in questa sezione, considereremo la nozione di confermabilità – in particolare la sua versione quantitativa (o bayesiana) – esclusivamente come criterio di demarcazione. Come abbiamo sostenuto riguardo alla verificabilità (sezione 2.3) e alla falsificabilità (cap. 3), anche la confermabilità verrà considerata (popperianamente) tracciare una linea di separazione tra enunciati scientifici (empiricamente controllabili) ed enunciati non scientifici o metafisici (non controllabili empiricamente) all’*interno* della classe degli enunciati che sono significanti, secondo il criterio logico-semantico di significato **CSS** (vedi sezione 2.2).

Così, in analogia con le formulazioni di **CVD** e **CFD**, possiamo formulare, in generale, il *criterio confermabilista di demarcazione* (**CCD**), come segue

> **CCD**. Un enunciato (non analitico) è scientifico (cioè empiricamente controllabile) se e solo se (i) soddisfa **CSS** e (ii) è, in linea di principio, confermabile o sconfermabile attraverso l'esperienza (osservazione); altrimenti l'enunciato è non scientifico (metafisico).

Ora, facendo uso delle definizioni quantitative di «conferma», di «sconferma» e «neutralità», introdotte alla fine della sezione 4.1, possiamo riformulare **CCD**, in termini rigorosamente bayesiani, come segue.

> **CCD'**. Un enunciato (non analitico) H è scientifico (cioè empiricamente controllabile) se, (i) soddisfa **CSS** e (ii) è, in linea di principio, possibile un'evidenza osservativa E, tale che $pr(H \mid E) > pr(H)$ oppure $pr(H \mid E) < pr(H)$;
> H è, invece, non scientifico o metafisico (cioè non controllabile empiricamente) se, per ogni possibile evidenza osservativa E, $pr(H \mid E) = pr(H)$ (cioè se ogni possibile evidenza osservativa è *neutrale* rispetto ad H)
> (ove per "evidenza osservativa possibile" si deve intendere evidenza osservativa o osservazione *empiricamente* (*fisicamente*) possibile, nel senso specificato nell'*Osservazione* 2.2).

In breve, per **CCD'**, un'ipotesi è scientifica, se la sua probabilità iniziale può essere, in linea di principio, *incrementata* o *decrementata* attraverso l'evidenza osservativa (o sperimentale); ed è, invece, non scientifica (o metafisica) se nessuna possibile evidenza osservativa (o sperimentale) è in grado di modificare la sua probabilità iniziale.

Possiamo ora considerare se **CCD'** soddisfa entrambe le condizioni del requisito di adeguatezza materiale introdotto nell'Introduzione, ed è, quindi, esente, dalla critica rivolta a **CVD** e **CFD** di escludere enunciati genuinamente scientifici, senza riuscire a escludere enunciati palesemente non scientifici (metafisici); e che costituisce, come abbiamo visto, il problema fondamentale della demarcazione.

Nelle sezioni 4.2.1 – 4.2.4 mostreremo che **CCD'** è sufficientemente ampio da includere tutte e quattro le classi di enunciati genuinamente scientifici considerate nelle sezioni 2.3.1–2.3.4 e 3.1.1–3.1.4; e nella sezione 4.2.5 mostreremo che è sufficientemente stretto da escludere tutti gli enunciati chiaramente metafisici (non scientifici).

4.2.1. La prima classe di enunciati, considerata nelle sezioni 2.3.1 e 3.1.1, è quella degli enunciati *generali illimitati con quantificazione uniforme* (*universale* ed *esistenziale*), come, per esempio, (1) $(\forall x)P(x)$ e (2) $(\exists x)P(x)$. Abbiamo visto che gli enunciati di questa classe sono *semidecidibili*: quelli universali, essendo unilateralmente falsificabili (cioè falsificabili ma non verificabili), e quelli esistenziali, unilateralmente verificabili (cioè verificabili ma non falsificabili).

Nella teoria bayesiana della conferma, come abbiamo visto nella sezione 4.1, la falsificazione corrisponde al valore di probabilità minimale 0 e la verificazione al valore di probabilità massimale 1; per cui falsificazione e verificazione rappresentano, rispettivamente, casi limite di conferma e sconferma. In particolare, per il Teorema di Bayes, un'evidenza osservativa E conferisce a un'ipotesi H il grado di sconferma $pr(H \mid E) = 0$ (falsificazione), se $H \Rightarrow E$ e $pr(E) = 0$; mentre E conferisce ad H il grado di conferma $pr(H \mid E) = 1$ (verificazione), se $E \Rightarrow H$ e $pr(E) = 1$. Ne segue che, dato un enunciato universale illimitato $(\forall x)P(x)$, da $(\forall x)P(x) \Rightarrow P(t)$ e $pr(P(t)) = 0$, si deriva $pr(((\forall x)P(x) \mid P(t))) = 0$; ma poiché $(\forall x)P(x)$ non è deducibile da alcun insieme finito di enunciati osservativi $P(t_1)$, …, $P(t_n)$, cioè $(P(t_1) \wedge, \ldots, \wedge\ P(t_n)) \not\Rightarrow (\forall x)P(x)$, anche se $pr(P(t_1) \wedge, \ldots, \wedge\ P(t_n)) = 1$, non è possibile derivare $pr((\forall x)P(x) \mid P(t_1) \wedge, \ldots, \wedge\ P(t_n))=1$ (a meno che $pr(((\forall x P(x)) = 1)$. Così un enunciato universale illimitato può essere sconfermato al massimo grado (cioè, falsificato) anche da una sola evidenza osservativa contraria, mentre non può essere confermato al massimo grado (cioè verificato) da alcun insieme finito di evidenze osservative. Al contrario, dato un enunciato esistenziale illimitato $(\exists x)P(x)$, da $P(t) \Rightarrow (\exists x)P(x)$ e $pr(P(t)) = 1$ si deriva $pr((\exists x)P(t) \mid P(t)) = 1$;

ma, poiché da $(\exists x)P(x)$ non è deducibile alcun insieme finito di enunciati osservativi $P(t_1), \ldots, P(t_n)$, cioè $(\exists x)P(x) \nRightarrow (P(t_1) \wedge, \ldots, \wedge P(t_n))$, da $pr(P(t_1) \wedge, \ldots, \wedge P(t_n)) = 0$ non si deriva $pr((\exists x)P(x) \mid P(t_1) \wedge, \ldots, \wedge P(t_n)) = 0$. Così, un enunciato esistenziale illimitato può essere confermato al massimo grado (cioè, verificato) anche sulla base di una singola evidenza osservativa, ma non può essere sconfermato al massimo grado (cioè, falsificato) da alcun insieme finito di evidenze osservative contrarie (si veda, tuttavia, l'*Osservazione* 4.2).

Il Teorema di Bayes permette, tuttavia, di valutare sia il *grado di conferma* che l'evidenza positiva disponibile E_1 conferisce a un'ipotesi universale, unilateralmente falsificabile, H_1, sia il *grado di sconferma*, che l'evidenza contraria disponibile E_2 conferisce ad un'ipotesi esistenziale, unilateralmente verificabile, H_2. Poiché E_1 conferma H_1 se e solo se $pr(H_1 \mid E_1) > pr(H_1)$, il grado di conferma fornito ad H_1, dall'evidenza (positiva) E_1, sarà tanto maggiore, quanto maggiore sarà il valore di $pr(H_1 \mid E_1) - pr(H_1)$; e, poiché E_2 sconferma H_2 se e solo se $pr(H_2 \mid E_2) < pr(H_2)$, il grado di sconferma conferito ad H_2, dalla evidenza (contraria) E_2, sarà tanto maggiore, quanto maggiore sarà il valore di $pr(H_2) - pr(H_2 \mid E_2)$.

In questo modo, la teoria bayesiana della conferma cattura, tra l'altro, la nozione popperiana di *grado di corroborazione* (vedi cap. 3) traducendola nei termini rigorosamente probabilistici dei gradi di conferma.

Il fatto che la teoria bayesiana riesce a determinare i gradi di conferma e di sconferma di enunciati che non sono, in linea di principio, né verificabili, né falsificabili conclusivamente, è rilevante per la soluzione del problema degli enunciati completamente indecidibili, che verranno considerati nelle successive tre sezioni.

4.2.2. La seconda classe di enunciati, considerata nelle sezioni 2.3.2 e 3.1.2, è quella degli *enunciati generali illimitati con quantificazione mista* di forma, per esempio, (3) $(\forall x)(\exists x)R(x, y)$ e (4) $(\exists x)(\forall x)R(x, y)$.

Nella sezione 2.3.2 abbiamo visto che questi enunciati sono completamente indecidibili: cioè, in linea di principio, non verificabili e non falsificabili in modo conclusivo. La ragione è che gli enunciati generali illimitati di forma (3) e (4) sono costituiti da combinazioni di quantificatori universali ed esistenziali. E abbiamo visto che un enunciato illimitato universalmente quantificato non è verificabile, in quanto non è implicato da alcun insieme o congiunzione finita di enunciati osservativi, e che un enunciato illimitato esistenzialmente quantificato non è falsificabile in quanto non implica alcun insieme o congiunzione finita di enunciati osservativi. Di conseguenza, un enunciato generale illimitato di forma (3) o (4), per la presenza del quantificatore universale non può essere implicato da alcun insieme di enunciati osservativi e, quindi, non può essere verificato; e per la presenza del quantificatore esistenziale non può implicare alcun insieme finito di enunciati osservativi e, quindi, non può essere falsificato.

Ma, se tra questi enunciati e gli enunciati base, che costituiscono i resoconti osservativi, non sussistono relazioni logiche strette di implicazione (conseguenza o deducibilità), possono sussistere, tuttavia, relazioni di dipendenza probabilistica, sufficienti a stabilire il grado di probabilità (o di conferma) conferito a un enunciato di forma (3) o (4) dall'evidenza osservativa disponibile. Così, data un'ipotesi H di forma (3) o (4), e un'evidenza osservativa E, il valore della probabilità finale di H rispetto a E sarà $0 < pr(H \mid E) < 1$, ove i valori estremi 1 (verificato) e 0 (falsificato) sono esclusi.

Pertanto, gli enunciati generali illimitati, con quantificazione mista, potendo essere, in linea di principio, confermati o sconfermati attraverso l'esperienza (osservazione), sono, in base a **CCD'**, empiricamente controllabili e, quindi, scientifici.

4.2.3. La terza importante classe di enunciati, considerata nelle sezioni 2.3.3 e 3.1.3, è quella degli *enunciati di probabilità statistica* di forma (5) $pr(A \mid B) = r$.

Nella sezione 2.3.3 abbiamo visto che nessun resoconto osservativo, descrivente le *frequenze relative* effettivamente rilevate di eventi di tipo A entro una classe-campione finita di eventi di tipo B può verificare o falsificare in modo conclusivo un enunciato di forma (5). La ragione è che nessun resoconto *finito* di questo tipo può implicare o essere implicato logicamente da un enunciato di forma (5).

Sempre nella sezione 2.3.3, abbiamo accennato, tuttavia, alla possibilità di controllare, attraverso la teoria bayesiana della conferma, gli enunciati di probabilità statistica come (5), calcolando, per mezzo del Teorema di Bayes, la probabilità epistemica (o grado di conferma) di un'ipotesi statistica come (5), sulla base dell'evidenza empirica disponibile, costituita dai resoconti osservativi delle frequenze relative effettivamente osservate: si tratta, cioè, di valutare la *probabilità epistemica* di una *probabilità statistica*. Cercheremo di chiarire brevemente questa idea.

A tal fine, indichiamo con *pre* la (funzione di) *probabilità epistemica*, con *prs* la (funzione di) *probabilità statistica* e con *RFR* il resoconto delle frequenze relative effettivamente rilevate (di eventi di tipo A entro la classe-campione di eventi osservati di tipo B), che costituisce l'evidenza osservativa di cui si dispone. Allora, il valore da stimare sarà $pre(prs(A \mid B) = r \mid RFR)$, che rappresenta la probabilità epistemica dell'ipotesi di probabilità statistica $prs(A \mid B) = r$, relativamente all'evidenza osservativa disponibile *RFR*, e che è calcolabile, attraverso la seguente formula, ottenuta dal Teorema di Bayes, sostituendo nella formula (**TB)**, introdotta nella sezione 4.1, ogni occorrenza di H con $prs(A \mid B) = r$ e ogni occorrenza di E con *RFR*:

$$pre(prs(A \mid B) = r \mid RFS) = \frac{pre(prs(A \mid B) = r) \cdot pre(RFS \mid prs(A \mid B) = r)}{pre(RFS)}$$

Questa formula – unitamente a certe altre assunzioni (o condizioni), come il Principio Principale di Lewis, che tralasciamo per semplicità – permette di determinare la probabilità epistemica della probabilità statistica espressa dall'ipotesi $prs(A \mid B) = r$, e, quindi, il grado di conferma o sconferma conferito a tale ipotesi, dalla evidenza osservativa disponibile (per una analisi più approfondita si veda Howson e Urbach (1989, capp. VIII – X e XIII-XIV; Festa, 1996, cap. 2, §. 11).

Pertanto gli enunciati di probabilità statistica, essendo, in linea di principio, confermabili o sconfermabili attraverso l'esperienza (osservazione) sono, in base a **CCD'** empiricamente controllabili, e, quindi, scientifici.

Osservazione 4.1. Nella sezione 3.1.3 abbiamo visto come Popper – nel tentativo di aggirare la seria difficoltà che l'infalsificabilità degli enunciati di probabilità statistica pongono al suo criterio di falsificabilità – abbia suggerito una disinvolta soluzione metodologica al problema logico della (non) falsificabilità degli enunciati statistici, sostenendo che tali enunciati, pur essendo *in linea di principio* *infalsificabili*, possono essere trattati come *falsificabili in pratica*, nel senso che uno scienziato può *decidere*, con l'ausilio di alcune regole convenzionali, se *accettare* un'ipotesi statistica come "confermata [corroborata] in pratica" o se rifiutarla come "falsificata in pratica". Ma, come abbiamo visto, questa soluzione presenta due gravi difetti. In primo luogo, come ha sottolineato Lakatos, confonde i concetti di «conferma [corroborazione]» e «falsificazione» con i concetti di «accettazione» e «rifiuto», rispettivamente. In secondo luogo, come hanno sottolineato Howson e Urbach, non riesce a fornire un criterio razionale oggettivo per l'accettazione e il rifiuto, facendo dipendere, in ultima analisi, la scelta da una decisione arbitraria e idiosincratica che fa appello sostanzialmente a ciò che Popper (1934/59, p. 64, nota 2) chiama "istinto del ricercatore".

Negli ultimi decenni sono state formulate alcune teorie bayesiane dell'accettazione, che, nonostante presentino vari problemi, sono almeno esenti da entrambi i difetti della soluzione popperiana.

Una di queste teorie, che appare particolarmente plausibile, almeno *prima facie*, è la *teoria (strettamente) probabilistica dell'accettazione*, che, pur definendo l'accettazione e il rifiuto in termini puramente probabilistici, differisce in modo rilevante dalla teoria della conferma.

Abbiamo visto che, per la teoria bayesiana della conferma, un'ipotesi H è confermata da una evidenza osservativa E se e solo se $pr(H \mid E) > pr(H)$; è sconfermata da E se e solo se $pr(H \mid E) < pr(H)$; e non è né confermata, né sconfermata da E se e solo se $pr(H \mid E) = pr(H)$. Può accadere, allora, che H risulti confermata da E anche se la sua probabilità finale $pr(H \mid E)$ rimane *bassa* o *molto bassa*, cioè $pr(H \mid E) < ½$; o che H risulti sconfermata da E anche se la sua probabilità finale rimane *alta* o *molto alta*, cioè $pr(H \mid E) > ½$. Al contrario, per la teoria probabilistica dell'accettazione, un'ipotesi H va accettata, sulla base dell'evidenza totale disponibile E, se e solo se $pr(H \mid E) > ½$; va rifiutata se e solo se $pr(H \mid E) < ½$; mentre va sospeso il giudizio riguardo ad H se e solo se $pr(H \mid E) = ½$. Poiché, per un teorema del calcolo della probabilità, la probabilità di un'ipotesi H è data da $pr(H) = 1 - pr(\neg H)$, possiamo dire che H va accettata, sulla base dell'evidenza totale disponibile E, se e solo se $pr(H \mid E) > pr(\neg H \mid E)$; va rifiutata se e solo se $pr(H \mid E) < pr(\neg H \mid E)$; mentre va sospeso il giudizio su H se e solo se $pr(H \mid E) = pr(\neg H \mid E)$.

A prima vista, questa teoria ha il vantaggio di fornire un criterio di accettazione oggettivo, particolarmente plausibile e semplice. Ma purtroppo produce due paradossi – noti come il *paradosso della lotteria* e il *paradosso della prefazione* – che mostrano che il criterio puramente probabilistico di accettazione porta ad accettare ipotesi contraddittorie, in violazione del fondamentale *requisito di coerenza*. E sembra che una tale conseguenza possa essere evitata solo a condizione di rinunciare al *requisito della congiunzione*, secondo cui, se due enunciati (ipotesi) H_1 e H_2 sono accettati, allora anche la loro congiunzione $H_1 \wedge H_2$ deve essere accettata, che per la sua intrinseca plausibilità è difficile abbandonare.

Oltre la teoria probabilistica, sono state proposte altre teorie bayesiane dell'accettazione, che fanno dipendere l'accettazione e il rifiuto delle ipotesi da obiettivi cognitivi diversi dal grado di probabilità, come la *verità*, la *verisimilitudine* o la *combinazione di probabilità e contenuto informativo* (che è particolarmente interessante, in quanto, facendo dipendere l'accettazione di una ipotesi, non solo dalla probabilità che ha di essere vera, ma anche dal suo contenuto informativo (empirico), cattura un aspetto importante della metodologia popperiana (vedi cap. 3)).

Non discuteremo qui queste teorie. Ci limitiamo solo a osservare che nessuna di esse sembra esente da limiti e difficoltà. La teoria basata sulla verità sembra presentare difficoltà analoghe alla teoria probabilistica. La teoria della verisimilitudine si basa su una nozione, come quella popperiana di *approssimazione alla verità*, che è stata criticata efficacemente da Miller (1974) e Tichý (1974); e non è chiaro se le recenti riformulazioni bayesiane della verisimilitudine siano completamente immuni da tale critica, e, nel caso, a quale costo. La teoria basata sulla combinazione di probabilità e contenuto informativo, nonostante la sua ragionevolezza epistemologica, sembra portare a risultati in conflitto, sia con le nostre intuizioni, che con la pratica scientifica (un'ampia trattazione delle teorie bayesiane dell'accettazione si trova in Festa (1996, cap. 7), che parteggia apertamente per la teoria della verisimilitudine).

Si deve, pertanto, concludere che non si dispone ancora di una teoria bayesiana dell'accettazione comparabile, per adeguatezza e efficacia, alla teoria bayesiana della conferma. Va, tuttavia, riconosciuto ai tentativi bayesiani il merito di aver cercato di fornire un criterio razionale oggettivo per l'accettazione e il rifiuto delle ipotesi, che evita ogni ricorso a decisioni arbitrarie e soggettive e ogni confusione tra accettazione e conferma e tra rifiuto e sconferma.

4.2.4. L'ultima classe di enunciati, considerata nelle sezioni 2.3.4 e 3.1.4, è quella degli *enunciati singolari su eventi passati* (in particolare, del passato remoto).

Abbiamo visto che questi enunciati sono, in linea di principio, completamente indecidibili: non possono essere verificati né falsificati, né *direttamente*, né *indirettamente*. Un enunciato (o ipotesi) singolare I intorno al passato, infatti, non può essere verificato o falsificato attraverso l'osservazione diretta, dal momento che il passato non è più accessibile epistemicamente. Inoltre, I non può essere verificato indirettamente, attraverso la verificazione dei resoconti osservativi E (descriventi le "tracce" presenti (o "fonti") dell'evento passato) derivabili da I con l'ausilio di un insieme Γ di leggi di copertura e ipotesi ausiliarie, cioè $(\Gamma \wedge I) \Rightarrow E$, poiché, come abbiamo visto nella sezione 2.3.4, da $(\Gamma \wedge I) \Rightarrow E$ e da una prova (della verità) di E (conclusione) non è derivabile una prova (della verità) della premessa $\Gamma \wedge I$; e non può essere neanche falsificato indirettamente, attraverso la falsificazione di un resoconto osservativo *E*, poiché, per la *tesi di Duhem*, come abbiamo visto nelle sezioni 2.3.4 e 3.1.4, da $(\Gamma \wedge I) \Rightarrow E$ e da una prova della falsità di *E* è derivabile, per *modus tollens*, la prova della falsità della premessa $\Gamma \wedge I$, *globalmente presa*, ma non una prova della falsità di una singola premessa Γ o *I*.

Così, verificazionismo e falsificazionismo non sono in grado di fornire una soluzione positiva del problema della controllabilità empirica degli enunciati intorno al passato.

Una soluzione positiva è fornita, invece, dalla teoria bayesiana della conferma, che provvede un metodo per la conferma o sconferma indiretta di tali enunciati. In particolare, la teoria bayesiana permette di valutare, in base al Teorema di Bayes, i differenti effetti prodotti dalla conferma o dalla sconferma (o falsificazione) di una teoria da parte di una data evidenza osservativa, sulle probabilità (e, quindi, sui gradi di conferma o di sconferma) delle singole ipotesi della teoria.

Cominciamo dal caso della sconferma delle ipotesi di una teoria. In questo caso, la soluzione bayesiana coincide con la soluzione del problema di Duhem, a cui abbiamo fatto cenno nella *Osservazione* 3.5.

La tesi di Duhem può essere riformulata in termini bayesiani come segue. Per semplicità si consideri una teoria *T* estremamente semplificata, costituita solo da due ipotesi *H* e *I* , cioè $T = H \wedge I$. Supponiamo che $T = H \wedge I$ implichi un'evidenza osservativa *E*, cioè $(H \wedge I) \Rightarrow E$; e che né *H*, né *I* singolarmente prese implichino *E*, cioè $H \not\Rightarrow E$ e $I \not\Rightarrow E$. Supponiamo, ora, che *E* venga falsificata da una evidenza osservativa (o sperimentale) contraria $E_c = \neg E$, per cui $pr(E) = 0$. Allora, come abbiamo visto nelle sezioni 4.1 e 4.2.1, da $(H \wedge I) \Rightarrow E$ e $pr(E) = 0$ segue $pr(H \wedge I \mid E) = 0$ (che corrisponde alla falsificazione dell'intera teoria $H \wedge I$ da parte dell'evidenza contraria E_c; ma, poiché $H \not\Rightarrow E$ e $\not\Rightarrow E$, non segue né $pr(H \mid E) = 0$, né $pr(I \mid E) = 0$ (cioè non segue la falsificazione delle ipotesi singolarmente prese). Questa è, in estrema sintesi, la formulazione bayesiana della tesi di Duhem.

Ma abbiamo visto che la teoria bayesiana della conferma è in grado di calcolare la probabilità finale $pr(H \mid E)$ di un'ipotesi H rispetto a un'evidenza E, anche quando tra H e E non sussiste alcuna relazione logica stretta, ma solo una relazione di dipendenza probabilistica. Così è possibile calcolare le probabilità finali $pr(H \mid E_c)$ di H rispetto all'evidenza contraria E_c e $pr(I \mid E_c)$ di I rispetto alla stessa evidenza contraria E_c, che, in base al Teorema di Bayes, sono date da:

$$\text{(i)}\ \ pr(H \mid E_c) = \frac{pr(H) \cdot pr(E_c \mid H)}{pr(E_c)}; \quad \text{(ii)}\ \ pr(I \mid E_c) = \frac{pr(I) \cdot pr(E_c \mid I)}{pr(E_c)}.$$

Tenuto conto dei valori di (i) e (ii), è possibile, allora, valutare i differenti effetti prodotti dalla evidenza contraria E_c sulle probabilità di H e I , comparando la probabilità finale $pr(H \mid E_c)$ di H rispetto ad E_c, fornita da (i), con la probabilità iniziale $pr(H)$ di H, e la probabilità finale $pr(I \mid E_c)$ di I rispetto a E_c, fornito da (ii), con la probabilità iniziale $pr(I)$ di I.

Possiamo, quindi, calcolare il *grado di decremento* della probabilità finale $pr(H| E_c)$ di H, data l'evidenza contraria E_c, rispetto alla probabilità iniziale $pr(H)$ di H, che è fornito da $pr(H) - pr(H \mid E_c)$; e il *grado di decremento* della probabilità finale $pr(I \mid E_c)$ di I, data la stessa evidenza contraria E_c, rispetto alla probabilità iniziale $pr(I)$ di I, che è fornito da $pr(I) - pr(I \mid E_c)$.

Se il valore di $pr(I) - pr(I \mid E_c)$ è *maggiore* del valore di $pr(H) - pr(H \mid E_c)$, cioè $(pr(I) - pr(I \mid E_c)) > (pr(H) - pr(H \mid E_c))$, allora vuol dire che la probabilità finale $pr(H| \mathrm{E}_c)$ è più vicina alla probabilità iniziale $pr(H)$ di quanto non sia la probabilità finale $pr(I \mid E_c)$ rispetto alla probabilità iniziale $pr(I)$. Se, invece, il valore di $pr(I) - pr(I \mid E_c)$ è *minore* del valore di $pr(H) - pr(H \mid E_c)$, cioè $(pr(I) - pr(I \mid E_c)) < (pr(H) - pr(H \mid E_c))$, vuol dire che la probabilità finale $pr(I \mid E_c)$ è più vicina alla probabilità iniziale $pr(I)$ di quanto non sia la probabilità finale $pr(H \mid E_c)$ rispetto alla probabilità iniziale $pr(H)$. Se, infine, il valore di $pr(I) - pr(I \mid E_c)$ è *uguale* al valore $pr(H) - pr(H \mid E_c)$, cioè $(pr(I) - pr(I \mid E_c)) = (pr(H) - pr(H \mid E_c))$, allora vuol dire che la distanza della probabilità finale $pr(I |E_c)$ dalla probabilità iniziale $pr(I)$ è uguale alla distanza della probabilità finale $pr(H \mid E_c)$ dalla probabilità iniziale $pr(H)$. Nel primo caso, diciamo che l'evidenza contraria E_c sconferma l'ipotesi I più di quanto sconferma l'ipotesi H; nel secondo caso, diciamo che E_c sconferma, invece, l'ipotesi H più di quanto sconferma l'ipotesi I; nel terzo caso, diciamo che E_c sconferma le ipotesi H e I nella stessa misura (si veda Dorling, 1979; e anche Howson e Urbach, 1989, cap. 4, §. h.3 e Festa, 1996, cap. 2, §. 11).

In modo analogo viene risolto il caso della conferma delle singole ipotesi della teoria. Supponiamo, che $T = H \wedge T$ venga confermata da una evidenza positiva (favorevole) E_p. Il grado di conferma che l'evidenza E_p conferisce alla teoria $H \wedge I$ è determinato dalla probabilità finale $pr(H \wedge I \mid E_p)$ di $H \wedge I$ rispetto all'evidenza favorevole E_p, che può essere calcolata attraverso il Teorema di Bayes

$$pr(H \wedge I \mid E_p) = \frac{pr(H \wedge I) \cdot pr(E_p \mid H \wedge I)}{pr(E_p)}.$$

Il Teorema di Bayes permette di calcolare anche le probabilità finali $pr(H \mid E_p)$ di H rispetto all'evidenza E_p e $pr(I \mid E_p)$ di I rispetto alla stessa evidenza E_p, che sono date da

$$\text{(iii) } pr(H \mid E_p) = \frac{pr(H) \cdot pr(E_p \mid H)}{pr(E_p)}; \quad \text{(iv) } pr(I \mid E_p) = \frac{pr(I) \cdot pr(E_p \mid I)}{pr(E_p)}.$$

Tenuto conto dei valori di (iii) e (iv), è possibile valutare – come nel caso della sconferma – i differenti effetti prodotti dall'evidenza favorevole E_p sulle probabilità di H e I, comparando la probabilità finale $pr(H \mid E_p)$ di H rispetto all'evidenza favorevole E_p, data da (iii), con la probabilità iniziale $pr(H)$ di H, e la probabilità finale $pr(I \mid E_p)$ di I rispetto alla stessa evidenza E_p, data da (iv), con la probabilità iniziale $pr(I)$ di I.

Possiamo, quindi, valutare *il grado di incremento* della probabilità finale $pr(H \mid E_p)$ di H, data l'evidenza positiva E_p, rispetto alla probabilità iniziale $pr(H)$ di H, che è fornito da $pr(H| E_p) - pr(H)$; e *il grado di incremento* della probabilità finale $pr(I \mid E_p)$ di I data E_p, rispetto alla probabilità iniziale $pr(I)$ di I, che è fornito da $pr(I \mid E_p) - pr(I)$.

Se il valore di $pr(I \mid E_p) - pr(I)$ risulta *maggiore* del valore di $pr(H| E_p) - pr(H)$, cioè $(pr(I \mid E_p) - pr(I)) > (pr(H| E_p) - pr(H))$, diciamo che l'evidenza osservativa E_p conferma l'ipotesi I più di quanto conferma l'ipotesi H. Se il valore di $pr(I \mid E_p) - pr(I)$ risulta *minore* del valore di $pr(H| E_p)$ - $pr(H)$, cioè $(pr(I \mid E_p) - pr(I)) < (pr(H| E_p) - pr(H))$, diciamo che l'evidenza E_p conferma H più di quanto conferma I. Se, infine, il valore di $pr(I \mid E_p) - pr(I)$ risulta *uguale* al valore di $pr(H| E_p) - pr(H)$, cioè $(pr(I \mid E_p) - pr(I)) = (pr(H| E_p) - pr(H))$, diciamo che E_p conferma I e H nella stessa misura.

In questo modo, il problema sollevato dalla *tesi di Duhem*, secondo cui "non [si] può mai sottoporre al controllo dell'esperienza un'ipotesi isolata, ma soltanto tutto un insieme di ipotesi" (Duhem 1906/14, p. 211), viene completamente risolto dalla teoria bayesiana della conferma: se la controllabilità empirica viene identificata con la confermabilità bayesiana, allora non solo le teorie prese come un tutto, ma anche le singole ipotesi delle teorie posso essere sottoposte a controllo empirico.

Questa possibilità costituisce anche una soluzione positiva del problema della controllabilità empirica indiretta degli enunciati (ipotesi) intorno al passato. Infatti, se I sta per un enunciato (ipotesi) intorno a un evento passato, E sta per un resoconto osservativo descrivente le "tracce" (o fonti) osservabili dell'evento passato, e H sta per una legge di copertura che connette l'evento passato descritto da I con le tracce descritte da E, allora la teoria bayesiana della conferma rende possibile, come si è visto, non solo confermare o sconfermare l'insieme $H \wedge I$, attraverso l'accertamento dell'evidenza empirica E, ma anche determinare se E conferma o sconferma l'ipotesi storica I in una misura maggiore, minore o uguale di quanto conferma o sconferma la legge di copertura H.

In questo modo, gli enunciati intorno al passato possono venire confermati o sconfermati indirettamente attraverso l'esperienza (osservazione). Per tanto, in base **CCD'**, sono empiricamente controllabili, e, quindi, scientifici.

4.2.5. Nelle sezioni 4.2.1 – 4.2.4 abbiamo visto che **CCD'** è *sufficientemente ampio* da includere tutte e quattro le classi di enunciati genuinamente scientifici considerate. Pertanto esso soddisfa la condizione (ii) del requisito di adeguatezza materiale introdotto nell'Introduzione.

In questa sezione, mostreremo che **CCD'** è anche *sufficientemente stretto* da escludere tutti gli enunciati chiaramente metafisici (non scientifici); e che, pertanto, soddisfa anche la condizione (i) del requisito di adeguatezza materiale.

Nella sezione 4.2 abbiamo visto che per **CCD'** un enunciato (non analitico) H è empiricamente non controllabile e, quindi, non scientifico (metafisico) se e solo se ogni evidenza osservativa è *neutrale* rispetto ad H: cioè se per ogni possibile evidenza osservativa (o sperimentale) E, $pr(H \mid E) = pr(H)$. In particolare, si può mostrare che tale criterio è anche in grado di escludere dalle teorie scientifiche tutti gli enunciati metafisici.

Nella sezione 3.2., abbiamo visto che, se a una teoria falsificabile T viene aggiunto un enunciato metafisico, la nuova teoria che si ottiene rimane ugualmente falsificabile e, quindi, scientifica secondo il criterio falsificabilista di demarcazione (**CFD**), nonostante contenga un enunciato chiaramente metafisico. Avevamo esemplificato questo aspetto del criterio di falsificabilità, considerando due casi: il primo, e più semplice, consisteva nell'aggiungere semplicemente a T un qualsiasi enunciato metafisico M; il secondo consisteva nell'aggiungere a T l'enunciato metafisico più complesso $O \wedge (M \rightarrow O)$, per evitare che il problema venisse aggirato con una modifica del requisito falsificabilista di scientificità dei sistemi teorici. Nel seguito considereremo, per semplicità, solo il primo caso dal momento che le conclusioni cui giungeremo al riguardo sono facilmente applicabili anche al caso più complesso. L'argomento utilizzato era il seguente.

Data una teoria T, se T è scientifica, allora, per il criterio di falsificabilità popperiano, deve essere falsificabile. Ma T è falsificabile se e solo se da T è deducibile un enunciato osservativo O, cioè $T \Rightarrow O$, in modo che dalla falsificazione (negazione) di O si possa derivare, per *modus tollens*, la falsificazione (negazione) di T. Ora, supponiamo di aggiungere a T un enunciato metafisico M (per esempio, un enunciato significante, ma, in linea di principio, del tutto indecidibile, come l'enunciato di Poincaré-Pap, vedi sezione 2.1), in modo da ottenere una nuova teoria $T' = T \wedge M$. Ora, per la monotonicità della relazione standard di deduzione, tutti gli enunciati deducibili da T saranno anche deducibili da T'. Di conseguenza, se un enunciato osservativo O è deducibile da T, allora sarà anche deducibile da $T \wedge M$: se $T \Rightarrow O$, allora $(T \wedge M) \Rightarrow O$. Ne segue che dalla falsificazione (negazione) di O è derivabile, per *modus tollens*, la falsificazione (negazione) di $T \wedge M$. Pertanto, anche la teoria $T' = T \wedge M$ risulta falsificabile e, quindi, scientifica per il criterio falsificabilista di demarcazione (**CFD**), nonostante contenga un enunciato chiaramente metafisico. Pertanto, **CFD** non è in grado di escludere gli enunciati metafisici dalle teorie scientifiche.

Si può facilmente vedere che il criterio bayesiano di confermabilità **CCD'** non presenta questo limite. Nella sezione 4.2.4 abbiamo visto che se una teoria viene globalmente falsificata (o sconfermata) da un'evidenza contraria, la teoria bayesiana della conferma permette di valutare, attraverso il Teorema di Bayes, il modo in cui le probabilità delle singole ipotesi costituenti la teoria vengono modificate dall'evidenza contraria: in particolare, permette di valutare il differente grado con cui le probabilità finali delle singole ipotesi, data l'evidenza contraria, *decrescono* rispetto alle loro probabilità iniziali. Ma se una teoria T', contenente un enunciato metafisico M, viene falsificata da una evidenza contraria $E_c = \neg O$, la probabilità di M non sarà modificata dall'evidenza contraria E_c, dal momento che qualsiasi evidenza contraria E_c sarà *neutrale* rispetto a M: cioè $pr(M \mid E_c) = pr(M)$. Lo stesso risultato si ottiene, ovviamente, nel caso in cui $T' = T \wedge M$ venga, invece, confermata da un'evidenza osservativa favorevole. Di conseguenza M verrà escluso come metafisico (non scientifico) in base a **CCD'**. Pertanto, **CCD'** risulta sufficientemente stretto da escludere tutti gli enunciati metafisici, anche quando ricorrono entro teorie scientifiche; soddisfacendo, così, la condizione (i) del requisito di adeguatezza materiale.

Possiamo, quindi, concludere che il criterio bayesiano di demarcazione (**CCD'**), fornito dalla teoria bayesiana della conferma, soddisfa entrambe le condizioni (i) e (ii) del requisito di adeguatezza materiale, e costituisce, perciò, un criterio materialmente adeguato.

Osservazione 4.2. Finora, discutendo della verificabilità, della falsificabilità e della confermabilità bayesiana, abbiamo assunto, come non problematica, la tesi che tutti gli *enunciati singolari osservativi* (o *enunciati protocollari* o *enunciati-base*), che descrivono l'evidenza osservativa, possono essere conclusivamente verificati o falsificati attraverso l'osservazione diretta (vedi *Osservazione* 2.3).

Ma sia Popper (1934/59, cap. V), che Carnap (1936-37 e 1950), hanno convincentemente argomentato che nessun enunciato empirico – a cominciare dagli enunciati osservativi base – è suscettibile di una prova *conclusiva* (*certa*, *assoluta*, *infallibile*, *irrivedibile*). In altri termini, non esiste alcuna base certa e infallibile per la conoscenza empirica e scientifica: tutti gli enunciati empirici, inclusi gli enunciati singolari osservativi (che descrivono dati osservativi), sono *rivedibili*, *ipotetici*, *fallibili*.

Se è così, allora è esclusa, in linea di principio, la possibilità di ottenere una prova conclusiva della verità (verificazione) o della falsità (falsificazione) di un qualsiasi enunciato, ipotesi o teoria empirica.

In termini della teoria bayesiana della conferma, dire che non esiste alcun enunciato osservativo certo equivale a escludere la possibilità di assegnare a una qualsiasi evidenza osservativa E il valore di probabilità massimale $pr(E) = 1$ (certo vero) o il valore di probabilità minimale $pr(E) = 0$ (certo falso). Come ha osservato Jeffrey (1992, p. 35), il grado di probabilità (o di credenza) attribuibile a una proposizione osservativa non può mai (o quasi mai) essere esattamente uguale 1 o esattamente uguale a 0, per il fatto che c'è sempre la possibilità, sia pure minima, di essere tratti in errore. Di conseguenza, la probabilità a priori (o iniziale) di una qualsiasi evidenza osservativa deve essere, sia pure di pochissimo, minore di 1 e maggiore di 0, cioè $0 < pr(E) < 1$. Ma, se la probabilità a priori di ogni evidenza osservativa E è $0 < pr(E) < 1$, allora, per il Teorema di Bayes, la probabilità finale $pr(H| E)$ di una qualsiasi ipotesi empirica H rispetto a una evidenza E, sarà $pr(H| E) < 1$, anche quando $E \Rightarrow H$; e $pr(H| E) > 0$, anche quando $H \Rightarrow \neg E$ (vedi sezioni 4.1). Di conseguenza, il valore della probabilità finale di una qualsiasi ipotesi empirica H rispetto a una qualsiasi evidenza osservativa E, sarà sempre $0 < pr(H | E) < 1$, con esclusione dei valori estremi 0 (certo falso) e 1 (certo vero).

Va osservato, che il calcolo delle probabilità – su cui si basa la teoria bayesiana della conferma – pur limitandosi, attraverso l'assioma A_2, ad assegnare il valore di probabilità massimale 1 agli enunciati analiticamente (necessariamente) veri e il valore di probabilità minimale 0 agli enunciati analiticamente (necessariamente) falsi, non esclude, tuttavia, la possibilità di assegnare tali valori estremi anche agli enunciati empirici. Ma, come vedremo nella sezione 4.3, questa possibilità può venire esclusa attraverso il *requisito di coerenza stretta*, introdotto per risolvere il problema della *probabilità nulla* delle ipotesi universali illimitate, sollevato da Popper.

La teoria bayesiana ha così il vantaggio di essere l'unica teoria razionale della conoscenza che può fare effettivamente a meno di una base o fondamento certo. In questo modo, il *fallibilismo* dell'approccio metodologico ipotetico-deduttivo e, in particolare, popperiano (vedi cap. 3), viene recuperato nella sua versione più radicale, entro l'approccio bayesiano, ed esplicato in termini di *probabilismo radicale*: la conoscenza empirica (scientifica) è esclusivamente e irriducibilmente *conoscenza probabile*, che non può mai attingere la certezza, a meno che non si tratti di ipotesi (o teorie) *logicamente incoerenti*; in quest'ultimo caso, però, non si tratterebbe di ipotesi empiriche, ma di ipotesi analiticamente (necessariamente) false, a cui, per l'assioma A_2 del calcolo delle probabilità, va assegnata probabilità 0 (per una breve introduzione storica al probabilismo, culminante nel bayesianesimo, si veda Morini, 2003).

4.3. Altri aspetti della metodologia bayesiana

Oltre a fornire un criterio di demarcazione adeguato, l'approccio bayesiano è in grado di spiegare le principali regole o principi metodologici dell'approccio ipotetico-deduttivo. Tra questi ricordiamo lo speciale valore evidenziale attribuito alle predizioni improbabili (come abbiamo visto nella sezione 4.1, commentando il rapporto inverso, stabilito dal Teorema di Bayes, tra i valori di $pr(H \mid E)$ e di $pr(E)$); la preferenza accordata alle ipotesi semplici; il rifiuto delle ipotesi *ad hoc*; l'idea che differenti insiemi di evidenze forniscono un supporto alle teorie più forte di quello fornito da evidenze più ristrette; e il riconoscimento della natura fallibile (incerta) e irriducibilmente ipotetica di tutta la conoscenza empirica, esplicata in termini probabilistici (vedi *Osservazione 4.2*).

Inoltre – come hanno sottolineato Curd e Cover (1998, pp. 549-550) – rispetto alla versione standard del metodo ipotetico-deduttivo, che concepisce questi ed altri principi come separati e non connessi, il bayesianesimo fornisce una spiegazione e una esplicazione *unificata* di tali principi, sulla base della logica della conferma incorporata nel Teorema di Bayes.

Va, infine, ricordato che l'approccio *quantitativo* bayesiano riesce a risolvere anche alcuni paradossi della teoria *qualitativa* della conferma, come il *paradosso dei corvi* di Hempel (1945) e, in particolare, il *paradosso degli smeraldi* di Goodman (1954). (Per i vantaggi dell'approccio bayesiano si veda, in particolare, Howson e Urbach, 1989, cap. IV; Earman, 1992, cap. III; e Festa 1996, capp. V e VI).

Di fatto, il potere esplicativo e i successi del bayesianesimo sono tali che ben pochi filosofi della scienza contemporanei hanno potuto sottrarsi alla sua influenza. Naturalmente non sono mancate le critiche. Ci limitiamo qui a menzionare quattro obiezioni principali.

La prima e più comune obiezione riguarda quello che è certamente il probema centrale del bayesianesimo: la possibilità di attribuire, in maniera *non arbitraria*, le