Published as *I problemi epistemologici del risk management* in: Daniele Pace (a cura di) **Economia del rischio. Antologia di scritti su rischio e decisione economica,** Giuffrè, Milano 2004.

Il caso sfortunato delle distribuzioni di probabilità non osservabili†

di NASSIM NICHOLAS TALEB* e AVITAL PILPEL**

1. Introduzione

Uno dei problemi più gravi per chi è esposto al rischio si ha quando non si ha neppure una vaga idea dei rischi in cui si è incorsi. Un problema ancora più grave si ha quando pur non avendo una vaga idea de tali rischi, ci si illude di conoscerli con esattezza. Semplicemente, è necessario avere una distribuzione di probabilità in base alla quale calcolare i rischi e definire la verosimiglianza/probabilità che alcuni eventi si verifichino.

Queste distribuzioni di probabilità non sono direttamente osservabili, il che rende discutibile qualsiasi calcolo del rischio in quanto basato sulla conoscenza di tali distribuzioni. Si è in possesso di sufficienti dati? Se la distribuzione è, per esempio, la classica Gaussiana a forma di campana, allora la risposta è si, e si può dire di disporre di sufficienti informazioni. Se, tuttavia, la distribuzione non appartiene ad una famiglia così ben definita, allora non si dispone di dati sufficienti. Ma come si può conoscere la distribuzione con cui si ha a che fare? *Ebbene, dagli stessi dati*. Se è necessaria una distribuzione di probabilità per valutare il comportamento futuro della distribuzione sulla base dei suoi risultati passati e se, allo stesso tempo, è necessario conoscere gli eventi passati per determinare la stessa distribuzione di probabilità, allora si è di fronte a un circolo vizioso, un problema in cui ogni dato rinvia a se stesso, simile a quello affrontato da Epimenide, l'uomo cretese che si domandava se gli abitanti di Creta fossero bugiardi oppure no. E questo problema autoreferenziale è solo il punto di partenza.

Cos'è una distribuzione di probabilità? In termini matematici, si tratta di una funzione con alcune proprietà definita su un dominio di "risultati possibili", X, che assegna dei valori a (alcuni)

† Questo articolo è la traduzione (a cura di Giacomo Nocera) di "*On the Very Unfortunate Problem of Not Observing Probabilità Distributions*", in corso di pubblicazione.
Gli autori ringraziano i partecipanti al Symposium on Chance Discovery dell'American Association of Artificial Intelligence di Cape Cod (novembre 2002), alla presentazione alla Stanford University (marzo 2003), all'Istituto per la Ricerca e gli Studi Assicurativi (Roma, aprile 2003) e alla ICBI Derivatives Conference di Barcellona (maggio 2003).
* Empirica LLC e Courant Institute of Mathematical Sciences, New York University.
** Columbia University.

sottoinsiemi di X. Una distribuzione di probabilità descrive una proprietà generale di un sistema: un dado è *regolare* se la sua distribuzione di probabilità consente di definire valori attesi che soddisfano i requisiti richiesti dall'operatore speranza matematica. In sostanza, è del tutto analogo alla descrizione matematica di altre proprietà del sistema (si potrebbe, ad esempio, pensare alla definizione della sua massa semplicemente assegnandogli un valore numerico di due chili).

La funzione di probabilità è (solitamente, ma non sempre) derivata dalla specifica condotta che il sistema ha manifestato in passato: l'osservazione dei lanci passati di un dado potrebbero giustificare l'affermazione che , in pratica, il dado è regolare, e quindi è descritto correttamente dalla funzione di probabilità di cui si è parlato in precedenza.

Solitamente, con le serie storiche si usano le osservazioni passate e si determinano le caratteristiche delle osservazioni future sulla base di ciò che si è osservato in passato. Può sembrare molto elegante o troppo semplicistico, tuttavia è certo che: le proprietà degli eventi futuri somigliano a quelle degli eventi passati, le proprietà osservate in passato sono sufficienti, ed è possibile avere un'idea precisa dell'ampiezza delle osservazioni passate di cui bisogna disporre per fare inferenza sulle osservazioni future.

Tuttavia vi sono delle notizie peggiori. Alcune distribuzioni cambiano col passare del tempo, quindi, indipendentemente dal numero di dati a disposizione, non è possibile fare inferenza sulle caratteristiche dei rischi dell'evento che si vuole valutare. È possibile che le proprietà con cui si ha a che fare siano instabili, oppure è possibile che diventino instabili perché le azioni che si compiono su di esse le modificano.

Allora, da cosa dipende la confusione sul "*risk management* scientifico" nelle scienze sociali, piene di equazioni, di dati ma prive di validità empirica (questi metodi falliscono regolarmente nella stima del rischio significativo) o intellettuale (il tema di cui si è parlato prima). Si sta per caso dimenticando qualcosa?

Un esempio. Si consideri l'affermazione: "è un evento dieci sigma", che si usa frequentemente in ambito stocastico quando il soggetto, di fronte a un evento sfavorevole, raro ed imprevisto, lo razionalizza considerandolo come la realizzazione di un processo aleatorio del quale i momenti sono noti, non considerando la possibilità di aver usato una distribuzione di probabilità sbagliata.

In queste scienze (soprattutto nel'ambito dell'Economia Politica), il *risk management* è afflitto dal seguente problema cruciale: *non si osservano le distribuzioni di probabilità, ma solo il risultato di generatori casuali.* Gran parte delle sofisticazioni nell'ambito della scienza della misurazione del rischio (a partire da Markowitz, 1952) si sono indirizzate verso i dettagli matematici ed econometrici del processo, invece di comprendere che i rischi connessi all'utilizzo di un'errata distribuzione di probabilità possono rivelarsi ancora più pericolosi rispetto a quelli riconducibili alla stessa distribuzione. Tutto ciò riporta alla mente la storia dell'ubriaco che cerca le chiavi smarrite sotto la lanterna, perché "lì c'è luce". Si potrebbe, ad esempio, considerare il fallimento dell'*hedge fund* Long Term Capital Management situato a Greenwich, nel

Connecticut[1]. I soci hanno spiegato tale fenomeno come il risultato di un "evento dieci sigma", che dovrebbe verificarsi una volta soltanto nel corso della vita dell'universo. Probabilmente sarebbe più convincente ritenere che abbiano utilizzato una distribuzione di probabilità sbagliata.

Nell'ambito della presente analisi è importante concentrare l'attenzione sugli eventi catastrofici, poiché sono quelli che determinano le conseguenze più rilevanti –quindi, indipendentemente dal fatto che la loro probabilità di verificarsi sia bassa (ipotizzando che sia tanto bassa quanto ritenuto dagli operatori) il loro impatto sul valore atteso è molto alto. Gli eventi catastrofici così definiti verranno indicati, di seguito, come *cigni neri*. Karl Popper ha osservato[2] che quando si perviene a generalizzazioni del tipo "tutti i cigni sono bianchi", è sufficiente che esista *un solo* cigno nero perché questa conclusione venga smentita. Peraltro, prima di trovare il cigno nero *nessuna informazione* sui cigni bianchi – indipendentemente dal fatto che tali informazioni si basino sull'osservazione di uno, cento, o un milione di cigni bianchi – può essere utile per stabilire se l'affermazione: "tutti i cigni sono bianchi" sia vera oppure no. Si può quindi ritenere che i *risk manager* affrontino situazioni di questo tipo. Non soltanto non sono in grado di prevedere un evento catastrofico, ma *nessuna informazione sul comportamento passato del mercato* consente di ridurre la loro ignoranza – attribuendo, ad esempio, delle probabilità significative all'evento "cigno nero". L'unica cosa che essi possono onestamente affermare circa un evento catastrofico è che "esso potrebbe verificarsi". E se questo, di fatto, *si* verifica, allora è possibile che *invalidi completamente tutte le precedenti conclusioni sull'operatore del valore atteso*, proprio come verrebbe smentita l'affermazione "tutti i cigni sono bianchi" se si osservasse un cigno nero. Ma allora sarebbe troppo tardi.

Chiaramente, la statistica matematica non è in grado di prevedere se eventi catastrofici di questo tipo si verificheranno in futuro: questa *ipotizza* che i risultati del processo che si osserva siano generati da una distribuzione di probabilità di un certo tipo (solitamente si fa riferimento ad una distribuzione Gaussiana). Tuttavia, non è in grado di spiegare come mai sia preferibile una distribuzione dall'andamento così "ben definito" piuttosto che le distribuzioni con una distribuzione "catastrofica", e neppure è in grado di suggerire il comportamento da tenere nel caso in cui vi sia il sospetto che le distribuzioni di probabilità possano cambiare inaspettatamente.

Ciò ci spinge a considerare l'epistemologia. Per epistemologia si intende il problema della teoria della conoscenza, sebbene, in termini più pratici,in questa sede si cerchi di rispondere alla seguente domanda: *cosa è possibile sapere sul futuro, date le osservazioni passate*? Secondo chi scrive esistono diversi motivi, filosofici e scientifici, per ritenere che, nelle scienze economiche e sociali, *non* sia possibile escludere la possibilità di osservare "cigni neri".

[1] Si veda Lowenstein (2000).

[2] Viene utilizzato il termine "osservato" e non "notato", in quanto già Aristotele aveva condotto considerazioni analoghe nell'ambito dell'analisi dei fatti importanti.

2. Tre tipologie di processo decisionale e il problema del *risk management*

Si supponga di voler valutare la veridicità di una certa proposizione P – "L'attuale presidente degli Stati Uniti d'America è George W. Bush, Jr."; "Al prossimo lancio della moneta uscirà testa"; "Su un pianeta che ruota intorno alla stella Tau Ceti ci sono forme di vita avanzate".

Nel primo caso, si può essere facilmente *sicuri* della veridicità della proposizione se si hanno i dati giusti per sapere chi è il presidente degli Stati Uniti. Se si dovessero scegliere un'azione o un comportamento sulla base della veridicità (o falsità) di questa proposizione – come, ad esempio, scegliere se rivolgersi al signor Bush chiamandolo "Signor Presidente" – ci si troverebbe di fronte a un *processo decisionale in condizioni di certezza*.

Nel secondo caso, invece, non è possibile essere certi della veridicità della proposizione, ma è possibile risalire alla *probabilità* che questa sia vera. In pratica, non esiste alcun metodo per sapere se al prossimo lancio della moneta si osserverà "testa" oppure "croce"; tuttavia, sotto certe condizioni, si può concludere che: p('testa') = p('croce') = 0.5. Se si dovesse scegliere un'azione sulla base della veridicità (o falsità) di questa proposizione – come, ad esempio, scegliere di accettare una scommessa in base alla quale la possibilità di osservare "testa" è 1:3 – ci si troverebbe di fronte ad un *processo decisionale in condizioni rischiose.*

Nel terzo caso, non solo non è possibile valutare la veridicità della proposizione, ma non è neppure possibile assegnare ad essa una qualsiasi probabilità significativa. Non solo non è possibile sapere se esistano forme di vita avanzate sulla stella Tau Ceti; ma non si è neppure a disposizione di una qualsiasi informazione che consenta di stimarne la probabilità. Se si dovesse scegliere un'azione in base all'esistenza di tali forme di vita, ci si troverebbe di fronte a un *processo decisionale in condizioni di incertezza*[3].

In economia è assai rilevante il caso in cui si debbano prendere delle decisioni sulla base dell'eventualità che si verifichino determinati eventi di carattere sociale, economico o politico – come, ad esempio, lo scoppio di una guerra. Molti sono convinti che, quando il futuro dipende non solo dall'universo fisico ma anche dalle azioni umane, non esistono leggi –neppure di carattere probabilistico – in grado di determinare il risultato; si è quindi in una perenne "condizione di incertezza"[4]. Come sosteneva Keynes (1937):

> Per conoscenza "incerta" … non intendo semplicemente distinguere ciò che è noto con certezza da ciò che è soltanto probabile. In questo senso il gioco della roulette non è soggetto a incertezza… Il senso in cui utilizzo il termine è quello in cui l'eventualità di una guerra in Europa è incerta, o il prezzo del rame e il tasso d'interesse fra vent'anni, o, ancora il processo

[3] Per quanto riguarda la prima distinzione tra rischio ed incertezza si veda Knight (1921), da cui la contrapposizione tra "rischio" di Knight e "incertezza" di Knight. Si veda, inoltre, Keynes (1937).

[4] La delicatezza del problema dell'incertezza, soprattutto nel caso di eventi futuri di tale fattispecie, ha portato Ramsey (1931), De Finetti (1937), e Savage (1954) a sviluppare una visione "personalistica" e "soggettiva" della probabilità, assolutamente indipendente da qualunque probabilità oggettiva o dalla sua assenza.

> di obsolescenza di una nuova invenzione... Su questi temi non esiste alcun fondamento scientifico sulla cui base formulare un qualsivoglia tipo di calcolo delle probabilità. Semplicemente, non sappiamo![5]

Certezza, rischio e incertezza non differiscono soltanto con riferimento alle probabilità (o nell'intervallo di probabilità) assegnate a P̲, ma anche alle strategie che ciascuno deve mettere in atto per prendere una decisione sotto tali differenti condizioni. Di solito, in una situazione di "certezza" si sceglie l'azione il cui risultato garantisce la maggiore utilità; in una situazione di "rischio" si sceglie l'azione il cui risultato garantisce la maggiore utilità *attesa*[6]. Per le situazioni di (totale) "incertezza" sono state proposte diverse strategie. La più famosa è la strategia *minmax*[7][8], ma ne esistono molte altre[9]. Queste strategie richiedono distribuzioni limitate. Nell'ipotesi di distribuzioni non limitate, la letteratura non è stata in grado di fornire risposte efficaci.

3. Il principale problema del *risk management*

L'impiego di una strategia adatta alle scelte in condizioni rischiose anche in quelle situazioni che potrebbero essere meglio descritte come situazioni di incertezza potrebbe avere degli effetti devastanti. Se un losco individuo all'angolo di una strada proponesse di giocare al gioco delle tre carte, nel giro di poco tempo si perderebbe tutto se si considerasse il gioco come una situazione rischiosa con p(vincita) = 1/3. Bisognerebbe invece prendere in considerazione l'eventualità che il gioco sia truccato e che le probabilità effettive di vittoria siano pressoché nulle. L'incertezza sul valore assunto, nell'intervallo [0, 1/3], dalla vera probabilità di vittoria, dovrebbe indurre ad utilizzare la strategia adatta in condizioni di incertezza (*minmax*), rifiutando la scommessa.

Chi scrive ritiene che sia proprio la pratica del *risk management* (definita come il monitoraggio dell'eventualità e dell'entità di risultati sfavorevoli) a condurre gli operatori a commettere errori di questo tipo. Di seguito si rileverà come, per diverse ragioni, i *risk manager* non possono

---

[5] Dire che "non sappiamo" non è necessariamente un'affermazione pessimistica. Shackle (1955) strutturò la sua teoria economica essenzialmente su questa "fondamentale ignoranza" – vale a dire l'incertezza del futuro. E' proprio questa "ignoranza" che permette all'uomo di effettuare delle scelte efficaci, che gli consente di *creare*, grazie al suo impegno, il proprio futuro sulla base dell'"ignoranza – in altre parole grazie al suo libero arbitrio. Se le regole della probabilità fossero applicate al futuro nella stessa maniera in cui sono applicate ai giochi d'azzardo, sostiene Shackle, si potrebbe soltanto *prevedere* passivamente quale, tra diversi predeterminati possibili futuri, si manifesterà.

[6] Oggigiorno queste idee appaiono alquanto tautologiche, in realtà non è affatto così. Sono stati von Neumann e Morgenstern (1944) a convincere il mondo, attraverso una rigorosa analisi matematica, che è possibile attribuire una funzione significativa di utilità attesa a differenti azioni, quando ci si trova ad operare una scelta in una condizione di rischio o di incertezza; e che massimizzare questa utilità attesa è una cosa assolutamente "razionale" (rispetto ad altri parametri). L'idea dell'"utilità attesa" si ritrova già in Bernoulli (1738) e Cramer (1728), ma per diverse ragioni non ne fu riconosciuta l'importanza.

[7] I primi ad elaborare nel dettaglio la strategia "*minimax*" sono stati von Neumann e Morgenstern (1944) e Wald (1950).

[8] Il principio minmax richiede che l'agente scelga (e quindi metta in pratica) l'azione che rende minima la massima perdita attesa. E' come se si valutassero i possibili risultati attesi in diversi scenari e si scegliesse l'azione che nello scenario peggiore garantisce l'utilità massima, *n.d.t.*.

[9] Come, ad esempio, la strategia di "*minimax regret*" di Savage.

evitare gli "eventi catastrofici". Successivamente si mostrerà come l'inevitabile probabilità di cigni neri si traduca, in molte circostanze, nell'inevitabilità che i *risk manager* affrontino situazioni future *essenzialmente incerte* nel senso di Knight, per le quali non è possibile assegnare alcuna probabilità significativa a certe possibili osservazioni future.

Peggioancora, questo vuol dire che non è possibile definire un limite inferiore (o superiore) all'insieme dei possibili esiti. La cosa peggiore in assoluto, tuttavia, è che mentre è spesso possibile, attraverso il campionamento oppure altre azioni, ridurre l'incertezza, in molti casi nessun tipo di informazione può aiutare i *risk manager* a ridurre i margini di incertezza.

Il principale problema del *risk management* è che, a causa delle proprietà generali dei generatori con cui i *risk manager* hanno a che fare, essi devono affrontare una situazione di sostanziale incertezza e non una situazione rischiosa.

In termini più formali: i *risk manager* guardano un insieme di stati del mondo[10] con probabilità cumulata superiore a un numero fissato arbitrariamente. Ciò implica che un *generatore* appartenente ad una *classe generale* (ad esempio, una distribuzione di probabilità nota come la Normale, la Binomiale, la Poisson ecc., oppure semplici istogrammi di frequenza) determina le osservazioni. Tale generatore ha specifici *parametri* (cioè una specifica media, una deviazione standard e dei momenti di ordine superiore) che, unitamente alle informazioni relative alla sua appartenenza ad una classe generale, determinano i valori della sua distribuzione. Una volta determinata la distribuzione, il *risk manager* potrà calcolare il "rischio" – cioè, la probabilità – che si verifichino gli stati del mondo ai quali è interessato.

Nella maggior parte dei casi più significativi, sia nell'ambito delle scienze "semplici" o "difficili", il generatore non è noto. In primo luogo, *non* esiste un modo indipendente per determinare i parametri –cioè, la media, la deviazione standard ecc. – a meno che non si tenti di *dedurli dal comportamento passato* del generatore. In secondo luogo, affinché si possano stimare tali parametri, è necessario *ipotizzare* che il generatore in questione *appartenga* a una certa classe generale: per esempio, si potrebbe ipotizzare che sia un generatore Normale, Poisson ecc. Occorre allora stimare, insieme, generatore e parametri.

In alcune circostanze, è corretto assumere che il generatore appartenga ad una certa classe e che la stima dei parametri sulla base dell'andamento passato sia affidabile. Questo è il caso, ad esempio, in cui dal lancio ripetuto di una moneta è possibile osservare la natura del generatore e accertare la limitatezza dei risultati.

In altri casi, invece, è corretto assumere che il generatore sia di un certo tipo, ma *non* è corretto assumere che i dati passati consentano di determinarne correttamente i momenti, indipendentemente dalla quantità di dati di cui si dispone.

[10] La nozione di "insieme di stati del mondo" (*state-space*) a cui si fa riferimento è quella introdotta da Arrow-Debreu nell'ambito dell'economia neoclassica.

In circostanze più difficili, non solo non è possibile ipotizzare i parametri del generatore, ma neppure la *classe generale* alla quale lo stesso generatore appartiene. In queste circostanze, naturalmente, non ha senso attribuire un valore ai parametri del generatore, dal momento che non si conoscono neppure i parametri a cui fare riferimento.

E' possibile affermare che la maggior parte delle circostanze in cui i *risk manager* operano sono talmente "gravi" che è impossibile determinare il generatore in base ai dati passati oppure stimarne i parametri in modo opportuno. Ciò significa che qualunque relazione tra i rischi che essi calcolano per i "cigni neri" ed i rischi *effettivi* di tali eventi può essere del tutto casuale. Ci si trova quindi in uno stato di incertezza: non solo non si è certi che $\underline{X}$ avvenga, ma non si può neppure stimare correttamente la probabilità $p(\underline{X})$. L'errore più grave compiuto dai *risk manager* è di confondere l'incertezza con il rischio, definendo, in modo del tutto ingiustificato, sia la classe del generatore sia i suoi parametri.

Di seguito, questo problema verrà analizzato con un approccio "Gedanken"[11]. Successivamente verrà analizzata la condotta ottima (se esiste) in condizioni di incertezza sul generatore.

4. Quattro simulazioni Monte Carlo con approccio "Gedanken"

Si consideri un generatore non osservabile di un processo stocastico. Associato ad uno spazio di probabilità, il generatore produce risultati osservabili. Cosa si può inferire da tali osservazioni sul generatore – e, successivamente, sulle osservazioni future? Cosa si può dedurre eventualmente sulla media, la varianza e i momenti di ordine superiore, oppure con quale probabilità i risultati futuri possono riprodurre quelli passati?

La risposta dipende, chiaramente, dalle proprietà del generatore. Madre Natura, in ogni caso, non concede la capacità d'osservare il generatore; il che è ancora più grave nel caso dei generatori che caratterizzano le scienze sociali (specie le scienze economiche).

Si considerino quattro casi. In tutti questi casi si osservano soltanto delle semplici *realizzazioni*, mentre il generatore non è osservabile. Si ipotizzi che i risultati osservati siano prodotti attraverso una simulazione Monte Carlo da un soggetto che si rifiuta di rivelare il programma, ma che offre campioni delle serie prodotte.

[11] "Gedanken" è una parola tedesca che significa "pensiero". Un esperimento Gedanken è un esperimento realizzato solo mentalmente. In fisica, il termine "esperimento Gedanken" è usato con riferimento a un esperimento poco pratico da effettuare, ma è utile da considerare perché può essere oggetto di ragionamenti teorici, *n.d.t.*

Tabella 1: I quattro esperimenti *Gedanken*

Gedanken	**Spazio di probabilità**	**Processo selezionato**	**Effetto**	**Commenti**
1	Limitato	Bernouilli	Convergenza Veloce	Caso "più semplice"
2	Illimitato	Gaussiano (generale)	Convergenza Semi-veloce	Caso "semplice"
3	Illimitato	Gaussiano (misto)	Convergenza lenta	Problemi risolvibili
4	Illimitato	Pareto-Levy Stabile	Nessuna convergenza	Nessuna soluzione nota

4.1 Il caso "ordinario" (2): il dado

La tipologia più semplice di processo aleatorio (talvolta definito come "*chance setup*") si ha quando tutte le possibili realizzazioni del processo sono limitate. Un semplice esempio è il lancio di un dado. Lo spazio delle probabilità consente soltanto risultati discreti, compresi tra 1 e 6 (inclusi).

L'effetto di ottenere errati momenti della distribuzione non è pericoloso.
In primo luogo, si noti che il generatore è *limitato*: il risultato non può essere maggiore di 6 o minore di 1. Gli errori commessi nella stima della media sono limitati, così come accade per la stima degli altri momenti (sebbene il margine di errore possa essere relativamente elevato per momenti di ordine superiore)[12].

In secondo luogo, la limitatezza del generatore implica che *non vi sono eventi estremi*. Non ci sono eventi rari, a bassa probabilità che nessun tipo di generatore non sia in grado di coprire e che abbiano comunque un effetto significativo sul valore dei momenti effettivi. Certamente, *non vi sono cigni neri* –eventi il cui risultato potrebbe vanificare le stime precedentemente effettuate sui momenti del generatore, indipendentemente dall'ampiezza del campione di dati di cui si dispone. In altre parole, è probabile che $E(\underline{X}_n)$ (la media osservata) sia vicina alla media ("effettiva") $E(\underline{X})$ poiché non esiste un evento raro, con probabilità 1 su 1.000.000, come quello che il valore osservato dopo il lancio di un dado sia 1.000.000: eventualità che farebbe aumentare $E(\underline{X})$ rispetto dal valore associato a un dado "regolare" all'incirca ad 1 ma che lo situerebbe al di fuori dei risultati osservati $x_1, x_2, \ldots x_n$, a meno di non essere estremamente fortunati.

[12] Si noti la differenza tra "illimitato" e "infinito". Finché i momenti esistono, ogni errore di stima ha valore finito, nel senso che non può superare un certo livello, indipendentemente da quanto la stima sia arbitraria. Il punto è che tale livello *non è a priori limitato* da nessun vincolo.

4.2. Il caso "ordinario" (2): la distribuzione Normale

Una situazione più complessa si ha quando lo spazio di probabilità non è limitato. Si consideri una distribuzione normale con funzione di densità f_2. In questo caso, esiste una certa probabilità >0 che il risultato sia arbitrariamente alto o basso; ad esempio, che esso sia >M o <*m* per ogni M, $m \in \mathbf{R}$.

Tuttavia, al crescere di M e al decrescere di m, la probabilità che il risultato sia >M o <m diventa velocemente piccola.

Sebbene i risultati siano illimitati, l'identificazione del valore epistemologico dei parametri è semplificata dalla nozione di "compattezza" usata in ambito economico da Samuelson[13].

Una distribuzione compatta o, in termini estesi, una "distribuzione con supporto compatto", soddisfa la seguente proprietà matematica: i momenti M[*n*] si riducono esponenzialmente rispetto al momento secondo[14].

La distribuzione Gaussiana, inoltre, possiede un'altra importante proprietà: può essere *caratterizzata completamente dai suoi primi due momenti*[15]. Tutti i momenti di ordine superiore n={3,4,…,∞} sono semplici multipli di M[1] e M[2]. Quindi conoscere la media e la varianza della distribuzione consente di ricavare i momenti di ordine superiore. Quest'argomento verrà analizzato ulteriormente nel seguito del lavoro. Si noti, incidentalmente, che la distribuzione Gaussiana rappresenta la distribuzione di massima entropia condizionatamente alla conoscenza della media e della varianza.

Si può notare inoltre che la funzione di densità Gaussiana per una variabile aleatoria *x* è rappresentata come una successione progressiva di ragione $e^{-\frac{(x-m)^2}{2\sigma^2}}$, quindi la densità diminuisce molto velocemente all'aumentare di x, il che si riflette nell'assottigliarsi graduale della coda della Gaussiana. Un'implicazione particolarmente interessante è che, utilizzando il Teorema di Bayes, è possibile determinare la probabilità che (*x-m*) sia inferiore a 3σ oppure a 4σ, condizionatamente al fatto che esso è maggiore di 2σ, rispettivamente:

---

[13] Si veda Samuelson (1952).

[14] Un *momento non centrale* è definito come $M[n] \equiv \int_{\Omega} x^n \phi(x)\, dx$.

[15] Si consideri una particella W in uno spazio bidimensionale W(*t*). Essa si muove secondo gli incrementi casuali ΔW negli intervalli di tempo Δ*t*. Alle scadenze *t*+Δ*t* si avrà $W(t+\Delta t) = W + \Delta W + 1/2\Delta W^2 + 1/6\Delta W^3 + 1/24\Delta W^4 + \ldots$ Quindi calcolando il valore atteso di ambo i membri dell'equazione: E[W(*t*+Δ*t*)] = W +M[1]+ M[2]/2 +M[3]]/6+ M[4]/24 ecc. Poiché i momenti dispari sono pari a 0 mentre quelli pari sono multipli del momento secondo, approssimando secondo la formula di Taylor arrestata al momento secondo M[2] è possibile cogliere la maggior parte delle informazioni disponibili dal sistema.

$$\frac{\int_2^3 \phi\left(\frac{x-m}{\sigma}\right)dx}{\int_{-\infty}^2 \phi\left(\frac{x-m}{\sigma}\right)dx} = 94\% \text{ e } \frac{\int_2^4 \phi\left(\frac{x-m}{\sigma}\right)dx}{\int_{-\infty}^2 \phi\left(\frac{x-m}{\sigma}\right)dx} = 99.8\%.$$

4.3. Il caso "semi-pessimistico": la distribuzione è "strana" e i suoi momenti esistono

Si consideri un altro caso di distribuzione non limitata: questa volta, però, si tratta di una combinazione lineare di una distribuzione "ordinaria" (come la distribuzione Normale) con una distribuzione "strana" (come la distribuzione di Poisson), che assegna probabilità molto basse a risultati rari –come, ad esempio, la probabilità di osservare un "cigno nero".

Per rendere l'analisi più concreta, si ipotizzi che il campionamento avvenga da due distribuzioni gaussiane. La probabilità che il campionamento avvenga da una Normale N_1 con media μ_1 e deviazione standard σ_1 è pari a π_1, mentre la probabilità che il campionamento avvenga da una Normale N_2 con media μ_2 e deviazione standard σ_2 è pari a $\pi_2=1-\pi_1$.

Si ipotizzi che la distribuzione N_1 rappresenti il sistema in condizioni "normali", poiché la probabilità π_1 è elevata, mentre N_2 rappresenti il sistema in condizioni "insolite". Si ipotizzi, inoltre, che $|\mu_1| << |\mu_2|$ e che $|\sigma_1| << |\sigma_2|$. La funzione di densità f_3 di questa distribuzione sarà una combinazione lineare della densità di N_1 e N_2. Anche la sua funzione generatrice dei momenti M_3 è data come media ponderata delle funzioni generatrici dei momenti, M_1 e M_2, della distribuzione "ordinaria" e di quella "strana", in base al noto Teorema di Feller (1971)[16]. Questo vuol dire che gli stessi momenti $(\mu_3, \sigma_3, ...)$ sono combinazioni lineari dei momenti delle due distribuzioni normali.

Se è vero che le proprietà di questo generatore e i risultati attesi sono meno "stabili" (in un senso che sarà definito più avanti) rispetto a quelli relativi ai casi precedenti, è vero anche che per questa funzione generatrice la media, la varianza e i momenti più alti esistono. Inoltre, tale distribuzione con il passare del tempo approssima una distribuzione Gaussiana, anche se ad una velocità sconosciuta.

Tuttavia, questa constatazione non è del tutto consolante quando σ_2 o m_2 sono molto grandi rispetto a σ_1 e m_1, come si è fin qui ipotizzato. In questo caso, infatti, occorre un campione di proporzione inversa rispetto a π_2 per approssimare i momenti veri. Quando π_2 è molto piccolo, ad esempio 1/1000, sono necessarie almeno 1000 osservazioni per cominciare a rilevare l'effetto di σ_2 e m_2 su tutti i momenti.

[16] La media e la deviazione standard di tale distribuzione sono, rispettivamente, $m = \pi_1\mu_1 + \pi_2\mu_2$ e $\sigma = \sqrt{\pi_1(m_1^2 + \sigma_1^2) + \pi_2(m_{12}^2 + \sigma_2^2) - (\pi_1 m_1 + \pi_2 m_2)^2}$.

4.4. Il caso "pessimistico": nessun generatore fissato

Si consideri il caso in cui il generatore stesso non è fissato, ma cambia continuamente e in modo imprevedibile; in questo caso x_1 indica il risultati di un generatore G_1 al tempo t_1, x_2 il risultati di un generatore G_2 nel periodo t_2, e così via. In questo caso, ovviamente, non si ha una sola funzione di densità oppure una sola funzione generatrice dei momenti, né è possibile assegnare un solo momento al generatore che è variabile.

Equivalentemente, è come se ciascun risultato fosse prodotto da un generatore che non ha momenti –con media non definita, con varianza infinita e così via. Un generatore di questo tipo ha funzione generatrice dei momenti M_4 e densità f_4 –cioè la distribuzione Pareto-Levy[17], che è uno caso particolare di distribuzione"L" Stabile (cioè stabile secondo Levy).

5. Le differenze tra generatori

Si supponga di osservare i risultati $x_1, x_2, x_3... x_n$ dei generatori di tipo (1)-(4) descritti in precedenza: dal lancio del dado fino alla distribuzione di Pareto–Levy. A quali conclusioni si può pervenire in base alle informazioni possedute su ciascun caso? Per rispondere a questa domanda occorre innanzitutto definire la relazione tra i momenti campionari ($E(X_n)$, $Var(X_n)$ ecc.) ed i momenti veri del generatore. Successivamente occorre individuare l'epistemologia che consente di individuare la rilevanza di tali relazioni per *conoscere* i veri momenti.

5.1. Il primo e il secondo caso

Nel primo e nel secondo caso i momenti del generatore (cioè, $E_1(X)$, $Var_1(X)$, $E_2(X)$, $Var_2(X)$, e i momenti di ordine superiore) sono facilmente ricavabili dall'osservazione dei risultati ottenuti.

Per esempio, il momento campionario primo –la media $E(X_n)=(x_1+x_2+\ldots+x_n)/n$– converge velocemente verso il momento vero $E_1(X)$ o $E_2(X)$ al crescere di n. Allo stesso modo la varianza relativa al campione $\{x_1\ldots x_n\}$, $Var(X_n)$, converge verso $Var_1(X)$ o $Var_2(X)$. Lo stesso accade per i momenti di ordine superiore.

Per verificare quanto detto precedentemente –la rapida convergenza dei momenti campionari verso i momenti veri- è sufficiente considerare il momento primo, cioè la media. Nel primo caso (il dado), le osservazioni sono limitate, cosicché si avrà certamente $\min(X)<x<\max(X)$. Nel secondo caso (la distribuzione Normale) le osservazioni non sono limitate, ma la loro probabilità diminuisce sempre più quanto più si discostano dalla media.

[17] Si veda Samorodnitsky e Taqqu (1994). E' interessante notare che la distribuzione Pareto-Levy è definita soltanto per mezzo della funzione caratteristica, non per la sua densità, che non può essere espressa correttamente in termini matematici, ma soltanto come l'inverso della trasformazione di Fourier.

In altre parole, $p_i(x)*x\rightarrow 0$ velocemente al crescere di x verso valori estremi, sia per il primo sia per il secondo generatore (cioè, per i =1, 2). Nel primo caso ciò dipende dal fatto che $p_1(x)=0$ per $x<\min(X)$ o $x>\max(X)$; nel secondo caso dal fatto che $p_2(x)$ decresce più velocemente rispetto alla deviazione di x dalla media.

Ciò implica che l'effetto dei valori estremi sulla media del generatore, $E_i(X)=\Sigma_x x*p_i(x)$, è trascurabile in entrambi i casi, quello delle osservazioni limitate (i=1) e quello della distribuzione Normale (i=2). Equivalentemente, si ha che $\Sigma_x x*p_i(x) \sim \Sigma_{x \text{ valore non estremo}}\, x*p_i(x)$ per entrambi i generatori considerati.

Si considerino adesso i dati che si osservano effettivamente. Nonostante i valori estremi del generatore con bassa probabilità di verificarsi (se esistono) *non* sono del tutto osservabili attraverso risultati x_1, $x_2\ldots$ x_n, la media "empirica" $E(X_n)=(x_1+x_2+\ldots x_n)/n$ tende ancora a convergere verso $\Sigma_{x \text{ valore non estremo}}\, x*p_i(x)$. Quest'ultimo, come già anticipato, non sarà molto diverso dal vero valore atteso $E_1(X)$ or $E_2(X)$. In altri termini, non è necessario aspettare che un evento estremo si manifesti, anche se esiste una certa probabilità che un evento di questo tipo abbia luogo, per poter disporre di una stima ragionevole del valore reale di $E_1(X)$ o di $E_2(X)$ attraverso la media campionaria $E(X_n)$.

Per motivi analoghi, la varianza $Var(X_n)$ convergerà velocemente verso $Var_1(X)$ oppure $Var_2(X)$, e lo stesso accade per momenti di ordine superiore, sebbene x_1, x_2, … x_n non includano nessuno dei valori estremi che potrebbero verificarsi –se ve ne è qualcuno.

5.2. Il caso "semi-pessimistico"

Si supponga che il generatore che ha prodotto i dati osservati– cioè i risultati x_1, x_2,… x_n – sia del terzo tipo, si è nella circostanza "semi-pessimistica" di una combinazione lineare tra la distribuzione Normale e quella di Poisson.

In questo caso i valori estremi del generatore *non* possono essere trascurati nella determinazione del momento del generatore. Quindi, anche se $p_3(x)\rightarrow 0$ per x che si discosta dalla media, la probabilità non diminuisce in modo "sufficientemente veloce" da rendere i valori estremi trascurabili. Cioè, $p_3(x)*x$ non $\rightarrow 0$ al divergere di x verso valori estremi.

In tali circostanze si avrà $E_3(X) = \Sigma_x p_3(x)*x \neq \Sigma_{x \text{ valore non estremo}}\, p_3(x)*x$. Quindi, finché gli eventi rari ed estremi non si verificano, la media "empirica" $E(X_n)$ tende a convergere verso $\Sigma_{x \text{ valore non estremo}}\, p_3(x)*x$ – che potrebbe essere molto diverso da $E_3(X) = \Sigma_x p_3(x)*x$.

In altre parole, gli eventi rari ed estremi devono *verificarsi effettivamente* prima che la media $E(X_n)$ sia prossima a $E_3(X)$. Lo stesso vale per il $Var(X_n)$ e il $Var_3(X)$ e per i momenti di ordine superiore.

Ciò si evince dal fatto che per tali generatori la trasformazione è molto più lenta.

Inoltre, prima che si verifichi effettivamente un "cigno nero", i le osservazioni relative al secondo generatore (Normale) potrebbero essere *indistinguibili* rispetto a quelle che derivano dal terzo generatore (Normale + Poisson). Le implicazioni di questa indistinguibilità verranno discusse più avanti.

5.3. Il caso "pessimistico"

Nel caso "pessimistico" è possibile che tali problemi non possano essere affrontati. Ciò non dipende dal tempo di convergenza dei momenti campionari, $E(X_n)$, $Var(X_n)$ ecc., verso i momenti "veri", $E_4(X)$, $Var_4(X)$ ecc. Semplicemente, in questo caso i momenti non esistono. Ciò significa che, indipendentemente dal numero di osservazioni di cui si dispone, si potranno determinare momenti campionari, $E(X_n)$, $Var(X_n)$ o momenti di ordine superiore, prossimi ai valori "veri" –dal momento che questi ultimi non esistono.

5.4. Il problema dell'inferenza induttiva ed il suo legame con le relazioni matematiche analizzate in precedenza

Fin qui sono stati semplicemente *descritti quattro generatori* e si è cercato di rilevare la relazione matematica che essi implicano tra il valore dei momenti stimati ed i momenti veri (se esistono).

Adesso occorre verificare il modo in cui tali proprietà influenzano il problema considerato inizialmente: si tratta di comprendere in quali circostanze è possibile utilizzare i dati relativi alle osservazioni prodotte dal generatore per stabilirne il tipo ed i suoi parametri, e quindi essere in grado di prevedere il rischio dei risultati futuri.

Sarebbe opportuno sottolineare che anche se i due problemi –la *relazione matematica* tra i veri momenti del generatore e i momenti campionari, da una parte, e l'abilità di *prevedere i risultati futuri* prodotti dal generatore, dall'altra- pur essendo strettamente correlati non sono tuttavia identici. Il primo è un problema puramente matematico. Il secondo, invece, è un problema epistemologico.

Non è possibile trarre delle conclusioni efficaci sul futuro considerando *semplicemente* un insieme ristretto e specifico di risultati, i "dati sperimentali". Nella moderna letteratura[18] si parla di un insieme di conoscenze (*corpus of knowledge*), per sottolineare che è necessario disporre di un certo insieme di "informazioni di background".

Per esempio, non è possibile stabilire, basandosi sull'osservazione di un milione di lanci di una *sola* moneta, la probabilità di osservare "testa" al prossimo lancio. I "dati" da soli, ad esempio, non consentono di escludere la possibilità che al prossimo lancio non si osserverà né "testa" né "croce", ma che la moneta esploderà come una bomba atomica. Nonostante la stretta relazione *matematica* tra i momenti campionari ed i momenti veri, se non si dispone del corretto insieme di

[18]Si vedano, ad esempio, Levi (1980) e Kyburg (1974).

“informazioni di background”, non sarà possibile trarre alcuna conclusione *epistemologica* relativa alla condotta futura del generatore sulla sola base delle osservazioni effettuate. La ragione per cui si escludono risultati del tipo “la moneta esploderà come una bomba nucleare”dipende, in molti casi, dal fatto che si dispone del giusto tipo di “informazioni di background” –come, per esempio, la conoscenza delle leggi della fisica.

D’altra parte, anche se il generatore fosse “pessimistico” come quello di Pareto-Levy, la mancanza di una relazione *matematica* tra i momenti campionari e quelli veri non escluderebbe – in teoria! – la possibilità di trarre una conclusione *epistemologica* relativa ai risultati futuri del generatore. Ad esempio, se per miracolo, fosse possibile comunicare con un angelo che sussurri nelle orecchie il prossimo risultato del generatore prima che abbia luogo, allora una parte delle “informazioni di background” *includerebbe* semplicemente il risultato del generatore e si potrebbe predire il risultato futuro.

Tuttavia, casi come questo non si verificano mai mentre, come si vedrà di seguito, solitamente l’informazione matematica è necessaria, ma *non* sufficiente, per giungere alle conclusioni epistemologiche alle quali si è interessati.

5.5. L’informazione di background implicita e le aspettative

Come già anticipato, l’aspetto più interessante è la risoluzione del problema epistemologico *dato un certo tipo di informazioni di backgound*, che rappresenta la situazione pratica in cui i *risk manager* devono mostrare quanto valgono. Si ipotizzi che siano date le informazioni di background:

1) i risultati provengono da un qualsiasi generatore casuale;

2) tale generatore casuale continuerà a produrre risultati nel futuro;

3) non esiste alcun sistema indipendente per stimare il tipo di generatore ed i suoi parametri, se non per mezzo dei dati relativi alle osservazioni passate;

4) il generatore potrebbe appartenere ad una qualunque delle tipologie trattate in precedenza.

Le prime tre ipotesi sul rischio non sono controverse. La quarta, invece, lo è.

Il problema “epistemologico” che si cerca di affrontare è il seguente: *se* l’informazione di background data è quella precedentemente descritta, *cosa* è possibile concludere, *se una conclusione esiste*, sui momenti del generatore (e quindi sul suo comportamento futuro) attraverso 1) l’osservazione del comportamento passato del generatore e 2) le informazioni di background? Dal punto di vista “pratico”, invece, ci si domanda: *in quali circostanze* il generatore può essere di quattro tipi, o quantomeno del tipo “pessimista”, tipo 3 o 4?

Si può affermare che:

1) Se il generatore *può essere* del tipo 3 o 4 ("semi pessimista" o "pessimista"), ciò è sufficiente per *invalidare* la capacità di trarre conclusioni sul suo andamento futuro in base a quello passato; in particolare, è impossibile assegnare una *qualsiasi probabilità specifica* ai risultati futuri, il che conduce ad una situazione di incertezza come quelle descritte nell'introduzione.

2) I *risk manager* tipicamente si confrontano con situazioni in cui il generatore *può essere* del tipo 3 o 4.

5.6. Il problema dell'inferenza induttiva: prima parte

Si consideri inizialmente la prima parte del problema, denominata del "se-allora": si tratta di definire in quali circostanze sia possibile (oppure no) concludere qualcosa sui risultati futuri del generatore *se* si conosce (oppure no) qualcosa sul generatore o sulla tipologia alla quale appartiene.

Vi sono due alternative. È possibile che una certa informazione sui momenti sia una *conseguenza diretta* della conoscenza già posseduta su di essi. Ad esempio, se si è a conoscenza del fatto che i risultati di un generatore sono compresi tra *a* e *b*, si sa che anche il momento primo è limitato. Non si tratta di un problema decisionale o di scelta: per essere coerenti dal punto di vista logico, *bisogna* accettare tutte le conseguenze di questo tipo che sono implicate direttamente dall'informazione di background sui momenti.

Il caso più complicato è quello dell'*induzione*. Anche quando (come si ipotizza solitamente) tutte le conclusioni che seguono dall'informazione di background sono note, può non essere determinabile un valore specifico per i momenti. In questi casi non bisogna fissare necessariamente un valore specifico per essi. Inoltre, si potrebbe concludere che, in alcune circostanze, si è *giustificati per induzione* ad assegnare un certo valore alla media del generatore (ad esempio, "3.5" nel caso del "dado regolare") e ai momenti di ordine superiore.

Di seguito il tema dell'induzione sarà affrontato in modo più esaustivo. Tuttavia, prima di intraprendere tale analisi è necessario riassumere quanto sin qui è stato detto.

Come evidenziato da Peirce, si tratta di un vero e proprio *problema decisionale* di carattere epistemologico. Si è a conoscenza dell'informazione di background sul generatore ("un dado di qualsiasi tipo è stato lanciato") ed i risultati osservati in passato ("i risultati sono stati 4, 4, 3, 2, 1"). Bisogna decidere se una qualsiasi conclusione sui momenti del generatore basata sui dati osservati sia giustificata oppure no (ad esempio, "il dado è *regolare*" o, più formalmente, "il momento primo del dado è 3.5").

Per risolvere questo problema decisionale, come accade per ogni problema decisionale, è necessario considerare l'*obiettivo* (o gli obiettivi) che si vuole raggiungere e le *opzioni* tra le quali è possibile scegliere. Scegliere correttamente vuol dire scegliere l'opzione che consente di raggiungere meglio l'obiettivo che ci si è prefissati. Gli obiettivi perseguiti dal decisore possono essere i più svariati, dalla possibilità di vincere una guerra nucleare alla scelta del miglior ristorante. L'obiettivo dell'inferenza induttiva è la *ricerca di nuove informazioni* e, al tempo stesso, la *fuga dall'errore*. Allo stesso modo, le opzioni disponibili possono essere diverse, dal lancio del missile *Trident II* alla possibilità di recarsi al ristorante. Nell'inferenza induttiva, le opzioni riguardano la possibilità di *aggiungere nuove conclusioni* alle opinioni dell'individuo –in questo caso, si tratta di conclusioni sui momenti del generatore casuale[19].

Questi due obiettivi sono contrastanti: più informazioni si accettano, più è probabile che si commettano errori. Il problema è la ricerca delle nuove informazioni che non incrementano eccessivamente il rischio di commettere errori una volta sommate a quelle precedentemente possedute. Il risultato dell'inferenza deduttiva –la soluzione del problema decisionale- consiste nell'*aggiungere alle opinioni individuali le informazioni che meglio consentono di raggiungere gli obiettivi prefissati.* Aggiungendo queste informazioni, l'inferenza induttiva è giustificata dalle situazioni specifiche.

Si noti che l'informazione nulla –che consiste nel "non aggiungere nessuna informazione"- è sempre disponibile. Se l'opzione ottima è l'informazione nulla, allora l'inferenza induttiva giustificata consiste nel *non* fare alcuna inferenza. In questo caso non si è giustificati a trarre alcuna conclusione sui momenti del generatore sulla base dell'informazione di background e delle osservazioni passate. Come si vedrà successivamente, questa è la situazione più frequente.

Si noti, inoltre, che il solo fatto di avere a che fare con un'elevata probabilità di successo (cioè, la probabilità di commettere errore è bassa) di per se *non* è sufficiente per effettuare una scelta. Si consideri una lotteria che vende milioni di biglietti: la probabilità che ciascun biglietto sia vincente è 1/1,000,000; tuttavia, riconoscere che tale probabilità è, di per sé, bassa potrebbe essere sufficiente per concludere che l'n-esimo biglietto *non* sarà vincente, il che condurrebbe all'assurda conclusione che *nessun* biglietto consentirà di vincere la lotteria.

Di seguito si cercherà di formalizzare e quantificare la situazione decisionale affrontata da un agente. A tale scopo utilizzeremo il sistema sviluppato da Levi. Esistono, tuttavia, altre formalizzazioni dell'espistemologia del processo decisionale; infatti, uno degli autori (Pilpel) sta studiando le differenze tra tali sistemi. Nelle situazioni affrontate dal *risk management* descritte in seguito qualsiasi sistema condurrà alla stessa (pessimistica) conclusione.

5.7. Generatori del primo e del secondo tipo

Si supponga che un angelo possa rivelare che: "il fenomeno appena analizzato, con esiti x_1, x_2, … x_n, è determinato o da un generatore limitato (tipo 1) oppure da un generatore che produce una distribuzione normale (tipo 2). In ogni caso, non verranno rivelati né la sua media, né la varianza, né i momenti di ordine superiore; questi verranno resi noti solo dai dati osservati". E' possibile determinarli? La risposta è positiva. Sulla base delle analisi matematiche descritte in precedenza, anche se *non* si conoscono i parametri del generatore, si è certi che al crescere di n, $E(X_n)$

[19] C.f.r. Levi *et al.*

converge rapidamente verso il valore vero E(*X*) (e lo stesso accade per i momenti di ordine superiore). Si sa inoltre che le stesse considerazioni varrebbero anche se il generatore fosse capace di produrre risultati estremi (come nel caso della distribuzione Normale) e non fosse stato possibile osservarli, come si è ipotizzato in questo caso: non si sta cercando di comprendere il tipo di previsione che è possibile effettuare se un evento raro è *già avvenuto*, ma si vuole comprendere come comportarsi se un evento di questo tipo *non* si è *ancora* verificato.

La situazione è "più rosea" di quanto possa sembrare. Non solo si è certi che i momenti campionari osservati –ad esempio, E(X_n)- convergono rapidamente verso i momenti veri, E(*X*) in questo caso. Sia nel caso limitato sia nel caso della distribuzione normale è possibile quantificare la probabilità che il momento osservato E(X_n) differisca da quello reale E(*X*) in misura superiore ad un certo livello.

Questo vuol dire che l'agente che deve stimare E(*X*) può determinare, abbastanza velocemente, solo la probabilità massima che la differenza tra E(X_n) e E(*X*) sia maggiore di un certo ε - e che, per ogni ε fissato, tale probabilità diminuisca rapidamente.

Come è stato detto in precedenza, l'induzione riguarda un *problema decisionale* sulla la relazione che intercorre tra il *rischio di errore* e l'*incremento del valore dell'informazione* che si consegue aggiungendo delle nuove informazioni alle opinioni individuali. In questo caso, il problema riguarda la possibilità che l'agente sia giustificato ad accettare l'ipotesi che E(*X*) *sia definito entro un certo intervallo*, $E(X) \in [E(X_n) - \varepsilon, E(X_n) + \varepsilon]$. Il problema cruciale consiste nel comprendere se il rischio di errore associato all'accettazione di tale conclusione sia un prezzo equo per l'incremento del valore dell'informazione che deriva dall'aggiunta di tale affermazione nell'insieme di opinioni dell'agente.

Questo è esattamente il principale problema di questo tipo di generatori. Poiché l'agente conosce il rischio di errore nel quale incorre aggiungendo questa ipotesi al proprio insieme di opinioni e poiché sa che tale rischio si riduce abbastanza velocemente, in molti casi è indotto a concludere che valga la pena di aggiungere un nuova informazione dal momento che il rischio di errore è basso e sono necessarie poche osservazioni $x_1, x_2, ..., x_n$ per accettare l'ipotesi che la differenza tra E(*X*) e E(X_n) sia inferiore ad un certo ε.

5.8. Generatori del terzo tipo – prima parte

Il problema è che, in molti casi, l'agente *non* sa se il generatore è del primo tipo o del secondo tipo. L'agente pertanto *fa un'ipotesi* giustificata dal fatto che essa consente di conseguire facilmente dei risultati apparentemente "corretti".

Si supponga, ad esempio, che fino ad ora la variazione giornaliera del prezzo di un titolo azionario sia stata compresa tra 0 e 10 punti. Ci sono ragioni per supporre che il prezzo di quel titolo non aumenterà o diminuirà di 1000 punti in futuro? Se si sapesse che il generatore che produce i movimenti di prezzo è normale, forse si. Ma spesso lo si ignora.

Si supponga, ad esempio, che un angelo affermi che "il fenomeno che stai osservando è generato dal generatore del terzo tipo. Questo è una combinazione di una distribuzione normale e di una distribuzione di Poisson che produce risultati estremi con probabilità molto basse. In ogni caso, non ti darò informazioni sulla media, sulla varianza né sui momenti di ordine superiore. Dovrai stimarli in base ai dati."

Osservando i dati, cosa si potrebbe dire a proposito della media, della varianza e degli altri momenti del generatore? Molto poco, in ogni caso – se non altro perché nessun evento catastrofico del tipo "cigno nero" di fatto *è* accaduto.

La ragione risiede nel fatto che per tale distribuzione il valore dei momenti dipende in gran parte dagli eventi rari ed improbabili, i cosiddetti "cigni neri", che dipendono dalla distribuzione estrema Poisson, e non dagli eventi ordinari e non catastrofici che si devono alla distribuzione Normale. Fino a quando nessun evento catastrofico accade, si conosce soltanto l'aspetto "negativo": per cui i momenti osservati E(X_n), Var(X_n) ecc. non rappresentano una buona approssimazione dei momenti veri E(X), Var(X), ecc. Ma questo è tutto ciò che si conosce e *non importa il numero di dati* (*non catastrofici*) *di cui si dispone*. Non si può dire nulla circa l'ordine dello scarto tra i momenti osservati e quelli veri fino a quando non si osservano eventi catastrofici.

In termini più formali sul piano epistemologico: come nel caso dei generatori del primo e del secondo tipo, si ipotizzi che John desideri individuare il momento primo E(X) in base a E(X_n)[20]. Come per i generatori del primo e del secondo tipo, il problema consiste nel comprendere se l'agente sia giustificato dal processo induttivo a concludere che E(X) differisca dal valore osservato E(X_n) non più di ε.

Inoltre, si tratta di un problema decisionale: il rischio di accettare erroneamente questa ipotesi è sufficientemente bilanciato dall'incremento del valore informativo che ne deriva? In questo caso, la risposta è solitamente negativa. Prima che si verifichi un evento catastrofico, l'agente *non sa* valutare il rischio che E(X) differisca rispetto a E(X_n) più di ε; e, in una situazione di questo tipo l'espansione induttiva (cioè, l'informazione aggiuntiva che deriva dal ritenere che E(X) differisca dal valore osservato E(X_n) non più di ε) non è giustificata.

In conclusione: anche se si è a *conoscenza* del fatto che il generatore è del terzo tipo (Normale + Poisson oppure Normale + Normale "strana" con media e varianza elevate), prima che un evento catastrofico accada non si può conoscere la differenza tra la media campionaria E(X_n) e E(X) (e lo stesso vale per i momenti di ordine superiore). Come viene evidenziato in dettaglio nell'appendice, questo è quanto accade *indipendentemente dal numero di osservazioni* verificatesi, finché si sa che nessun evento catastrofico si è verificato. Stimare i momenti del

[20] Come al solito le stesse considerazioni valgono, *mutatis mutandis*, per i momenti di ordine superiore.

generatore prima che si sia verificato un evento catastrofico è impossibile, indipendentemente dal numero di dati di cui si dispone.

In conclusione: anche se si è a *conoscenza* del fatto che il generatore è del terzo tipo (Normale + Poisson), prima che un evento catastrofico accada non si può conoscere la differenza tra la media campionaria E(X_n) e E(X), tra Var(X_n) e Var(X) e, più in generale, tra qualunque momento campionario e quello "vero". Prima che un evento di questo tipo si verifichi, è *inutile* cercare di dedurre dai dati passati il comportamento futuro di un qualsiasi sistema di questo tipo.

5.9. Generatori del quarto tipo – prima parte

Chiaramente, la situazione è ancora più difficile nel caso dei generatori del quarto tipo. Se un angelo afferma che un certo generatore è del quarto tipo (Pareto-Levy), si è certi che non esiste alcuna relazione tra i momenti campionari E(X_n), Var(X_n) ecc. e i momenti "reali" del generatore – poiché tali momenti non esistono.

Quindi, se si sta osservando un generatore di tipo 4 non si può giungere a nessuna conclusione, indipendentemente dal numero di dati osservati. Sia i generatori di tipo 3 sia quelli di tipo 4 sono portatori di "cattive notizie": prima ancora che un evento catastrofico accada (nel caso 3) o anche dopo (nel caso 4), è *inutile* cercare di dedurre dai dati passati il comportamento futuro di un qualsiasi sistema di questo tipo.

5.10. Generatori del terzo e del quarto tipo –seconda parte

Tuttavia le cose potrebbero essere addirittura peggiori rispetto ai casi fin qui esaminati. Si è appena detto che se un generatore è del tipo 1 o 2 è possibile fare affidamento sul fatto che i momenti campionari siano vicini ai momenti "reali". Si è anche detto che se un generatore è del tipo 3 o 4 i momenti campionari (almeno prima che un evento catastrofico o "cigno nero" accada) sono inutili nella stima dei momenti veri.

In tutti questi casi si è ipotizzato di conoscere la tipologia del generatore con cui si ha a che fare. Si supponga ora di *non* conoscere la classe cui il generatore appartiene e che si voglia verificare se i dati possano essere utili per stimare il generatore. In questo caso, l'uguaglianza matematica tra i momenti osservati e quelli campionari, *anche se è verificata* (cioè, anche se il generatore è del primo o del secondo tipo), potrebbe non essere sufficiente per formulare conclusioni epistemologiche sulla somiglianza tra gli eventi passati e quelli futuri. L'uguaglianza matematica è necessaria, ma non sufficiente.

Si consideri la situazione seguente. Si supponga che un angelo affermi che un certo generatore sia di tipo 2 (Normale) oppure di tipo 3 (una combinazione tra una distribuzione Normale e una Poisson). Si considerino i dati $x_1, x_2, \dots x_n$ che si hanno a disposizione. Finché nessun evento catastrofico "Poisson" è avvenuto, i dati sono *indistinguibili* e non è possibile stabilire se provengano da un generatore del secondo o del terzo tipo, poiché i risultati del generatore di tipo

3 potrebbero essere imputabili alla componente “Normale” della distribuzione. Non è quindi possibile stabilire da quale generatore siano stati prodotti i dati osservati.

Più in generale, si supponga che un angelo riveli che un certo risultato *potrebbe* essere dovuto ad un generatore del tipo 3 o del tipo 4, oltreché del tipo 1 o 2. Avendo a disposizione una certa quantità di dati, è possibile verificare la veridicità di questa affermazione prima che un evento “cigno nero” accada? No, finché non si verifica effettivamente un evento catastrofico a bassa probabilità, anche *se* i dati fossero generati da un generatore del tipo 3 o del tipo 4 sembrerebbero *indistinguibili* rispetto a quelli prodotti da un generatore del tipo 1 o del tipo 2.

Quindi, se *non si è a conoscenza* del fatto che il generatore *non* è del tipo 3 o 4 i dati a disposizione sono *inutili* per stabilire il comportamento futuro del generatore, come se fosse noto che il generatore sia del tipo 3 o 4. Ciò non dipende dal fatto che $E(X_n)$, $Var(X_n)$ e gli altri momenti *devono* essere lontani dai valori “reali” $E(X)$, $Var(X)$ ecc. (se questi esistono), ma dal fatto che non è possibile stabilire sulla base dei dati *se* questi si sono verificati prima di un evento catastrofico.

Allo stesso modo, se non si conoscono i momenti non è possibile definire le probabilità relative ai possibili esiti del generatore, che di fatto dipendono dai momenti. Quindi non è possibile neppure fare previsioni sul rischio relativo a un qualsiasi risultato futuro. Si tratta di una situazione di *scelta in condizioni di incertezza*.

In conclusione, perché l’inferenza sulle osservazioni future in base a quelle passate sia inutile non è necessario sapere che il generatore sia “pericoloso”: non è necessario che $E(X) \neq E(X_n)$ (lo stesso vale per i momenti di ordine superiore). E’ sufficiente sapere che il generatore *non* sia di questo tipo. In situazioni di incertezza come questa, un evento “cigno nero” potrebbe verificarsi in qualsiasi momento – e non è possibile stabilire il suo verificarsi fino a quando il fatto non sia realmente accaduto. L’uguaglianza matematica $E(X)= E(X_n)$ non è utile se non si *conosce in anticipo* che è valida anche quando un evento catastrofico si è verificato.

5.11. Possono esistere generatori di questo tipo?

L’intera analisi sarebbe rimasta su un piano puramente teorico se non accadesse che le situazioni con cui i *risk manager* si confrontano *potessero* comprendere anche generatori “cattivi” – e cioè a meno che non sia soddisfatta la quarta ipotesi epistemica analizzata in precedenza.

Come si è visto precedentemente, molti economisti escludono la quarta ipotesi. Eppure, sfortunatamente, nelle situazioni economiche generatori di questo tipo *possono* ricorrere. I sistemi fisici (come dice Mandelbrot) devono essere del tipo “benigno”– tipo 1 o 2, o, più specificamente, del tipo 1 (un generatore “limitato”). Le leggi della fisica limitano i valori da esse definiti –ad esempio, la quantità di energia di un sistema, la sua entropia e le altre caratteristiche fisiche non possono assumere valori al di fuori di un certo intervallo.

Nei sistemi fisici e sociali, quindi, è spesso possibile stabilire in anticipo, grazie a ragioni esterne puramente deduttive, che il "generatore" sia limitato e quindi di natura (relativamente) benigna; e si possano utilizzare i dati del passato per un'inferenza induttiva sul futuro, come già visto in precedenza.

In molti sistemi *finanziari* non è possibile pervenire a conclusioni di questo tipo. Nella maggior parte di questi sistemi alcuni eventi potrebbero causare perdite (o guadagni), in teoria, *illimitati.* Per convincersene è sufficiente riflettere su una semplice "opzione": la possibilità di perdere una quantità illimitata di danaro esiste e tale probabilità può rimanere sconosciuta.

Naturalmente, con queste considerazioni non si vuole dire che la morte sia in qualche modo "migliore" della perdita di un'enorme somma di denaro, o che guadagnare o perdere un'infinità (o una quantità molto, molto grande) di danaro sia fisicamente possibile. Si vuole dire, piuttosto, che nel caso di un sistema fisico, si è certi di poter descrivere il sistema con un generatore limitato (o, al peggio, con un supporto compatto), mentre lo stesso non accade per i sistemi finanziari.

6. La strategia consigliata in tali situazioni e il caso "Long Term Capital" riesaminato

La conclusione di questa rassegna epistemologica è la seguente: in tali situazioni, dunque, *ci si trova in una situazione sostanzialmente "incerta".*

Se si dovessero prendere delle decisioni in tali situazioni, sarebbe opportuno puntare su una strategia adatta alle condizioni di "incertezza". La *minmax* (o strategie simili) non funzionerebbe a causa dell'illimitatezza. "Imporsi" l'utilizzo un valore specifico della probabilità condurrebbe alla seguente conclusione: è inutile proteggersi dai rischi di certi risultati quando non è ragionevole attribuire ad essi una *qualsiasi* probabilità specifica.

Si noti, in particolare, che la ricorrente soluzione di fissare dei "margini di sicurezza" non funzionerà. Si supponga di essere disposti ad assumersi un rischio di fallimento con una probabilità pari a 1/1.000.000, e che –al fine di "coprire" il rischio– si effettuino delle operazioni che (secondo certi calcoli) hanno una probabilità pari a 1/mille miliardi di andare così male da portare al fallimento. Un "margine di sicurezza" così assurdo – un fattore pari a 1.000.000 – potrà aiutare il *risk manager* ad evitare il fallimento in tali situazioni?

La risposta è negativa. Assumere "misure di sicurezza" come quelle descritte costituisce un comportamento ragionevole se si *sa* che il generatore è "benigno", come ad esempio del tipo 1 o 2, e che *è* possibile formulare delle ipotesi sulla probabilità che certi eventi accadano in futuro utilizzando i parametri osservati per approssimare i parametri veri del generatore, pur non essendo completamente sicuri sui *veri valori* che i parametri dovrebbero avere. In altre parole, questo comportamento potrebbe funzionare nei casi in cui è possibile descrivere la situazione in cui ci si trova come una scelta in condizioni rischiose, sebbene non si conosca esattamente di quale rischio si tratti.

In una situazione in cui il generatore è del tipo 3 o 4, non si ha neppure una *vaga* idea della tipologia del rischio; non solo non si ha *idea* del tipo di rischio, ma non è possibile neppure assegnargli un *qualsiasi* valore. Accettare una scommessa che quota "mille miliardi a 1" contro l'eventualità di fallimento è inutile in una situazione del genere, poiché la stessa valutazione del rischio di una certa attività *come* una evento a probabilità mille miliardi a 1 è priva di significato e, analogamente, il " margine di sicurezza di un milione a uno" così calcolato non ha alcun significato nella realtà.

Non ci sono fondamenti oggettivi per dare credito a questa stima; pertanto, la cifra consolatoria di "mille miliardi a uno" ha soltanto un valore psicologico – come il verificarsi dell'evento "impossibile" 10-σ nel caso del "Long Term Capital". Non è come se un evento 10-σ si fosse effettivamente verificato. Piuttosto, la convinzione che esso sia un evento 10-σ si basava sulla *conclusione ingiustificata* che il generatore coinvolto fosse di tipo benigno.

Quindi, il *risk manager* non considera la possibilità che il generatore potesse essere del terzo o del quarto tipo, in qual caso gli eventi che *sarebbero* eventi 10-σ *se* il generatore fosse benigno accadrebbero di fatto molto più frequentemente.

In queste situazioni si può fare solo riscorso alla soluzione proposta da Popper: attendere il "cigno nero" ed assicurarsi di non essere distrutti.

7. Sintesi

In questo lavoro si è cercato di mostrare come il problema essenziale del *risk management* sia la tendenza a trattare le situazioni di decisioni in condizioni di incertezza come decisioni in condizioni di rischio. Lo si è fatto in quattro fasi.

Dapprima, è stato evidenziato che certi generatori aleatori non evidenziano un buon legame tra i momenti campionari e quelli veri. Tale problema è essenzialmente matematico.

In secondo luogo, è stato evidenziato che se le informazioni di partenza soddisfano alcune condizioni, quindi *se* tali generatori non sono scartati, allora la semplice possibilità che essi rappresentino il generatore con cui si ha a che fare rende vano qualsiasi tentativo di assegnare valori specifici ai momenti "reali" del generatore, a causa del problema dei "cigni neri" – cioè, la possibilità che si verifichino eventi rari ed estremi capaci di influenzare in modo significativo i momenti. Si tratta di un problema epistemologico.

In terzo luogo, è stato evidenziato che, di fatto, le situazioni con cui i *risk manager* si confrontano sono esattamente quelle situazioni in cui non si possono escludere tali generatori. Si tratta di un problema di teoria delle decisioni.

Successivamente, si è mostrato come le normali pratiche di "esclusione" – come, per esempio, la ritenzione dei rischi che appaiono "molto bassi" – non funzionano, poiché, implicitamente, assumono che la situazione sia quella in cui in primo luogo si prendono delle decisioni in condizioni rischiose. Persino le pratiche "comunemente" usate per fronteggiare le decisioni in condizioni di incertezza, *minmax*, *minmax of regret* [21]ecc., non sono efficaci, dal momento che i generatori "cattivi" non sono limitati.

Per concludere, si è mostrato che in tali situazioni la sola cosa che si possa fare è proteggersi dai "cigni neri" – ed ammettere che se ne ha solo una conoscenza limitata. Questa è l'unica strategia applicabile in tali circostanze.

[21] La strategia *minimax* consiste nel minimizzare la probabilità della perdita massima attesa. Nella variante *minimax of regret*, si cerca di minimizzare il rimpianto dovuto ad una scelta rivelatasi, a posteriori, errata. La principale differenza fra le due strategie afferisce all'approccio probabilistico sotteso: frequentista nel caso della prima strategia e soggettivistico nel caso della seconda. Infatti, nella prima strategia la perdita è misurata, sia *ex ante* che *ex post* in termini oggettivi mentre nella seconda l'entità attesa della medesima dipende dalla distribuzione di probabilità soggettivamente attesa e dai giudizi a priori che caratterizzano l'individuo mentre la misura *ex post* dipende dai peculiari profili cognitivi propri di ciascun individuo. [NdC]

Bibliografia

Bernoulli, D. *Specimen Theoriae novae de Mensura Sortis*, in *Commentarii Academiae Scientiarum Imperialis Petropolitanae* 5 (1738), St. Petersburg, 175-192.

Cramer, G. (1728), letter to Nikolaus Bernoulli, published by Bernoulli (1738).

De Finetti, Bruno. *La Prevision: ses lois logiques, ses Sources Subjectives*. In *Annales de l'Institut Henri Poincare*, 7 (1937), 1-68.

Feller, W. (1971). *An Introduction to Probability Theory and Its Application* (Vol. II; 2nd ed.), New York: Wiley.

Keynes, John Maynard. *The General Theory*. In *Quarterly Journal of Economics*, Vol. LI, 1937, 209-233.

Knight, Frank, 1921, *Risk, Uncertainty and Profit*, Harper Torchbook Edition, New York: Harper and Row, 1965.

Levi, Isaac, 1980, *The Enterprise of Knowledge,* Cambridge, MA: MIT Press.

Luce, R. Duncan and Raiffa, Howard. 1957. *Games and Decisions: Introduction and Critical Survey*. New York: John Wiley and Sons.

Markowitz, Harry, 1952, Portfolio Selection, *Journal of Finance* 7: 77-91.

Ramsey, Frank P. *Truth and Probability*. In *The Foundations of Mathematics and Other Logical Essays*. R. B. Braithwaite (ed.) London: Routledge and Kegan Paul,, 1931.

Samuelson, P. A., 1983 (1947), *Foundations of Economic Analysis,* Cambridge, MA: Harvard University Press.

Samorodnitsky, Gennady, and Murad S. Taqqu, 1974, *Stable Non-Gaussian Random Processes: Stochastic Models with Infinite Variance.* New York: Chapman & Hull.

Savage, Leonard J. *The Foundations of Statistics*. New York: John Wiley and Sons, 1954.

Shackle, G. L. S. *Uncertainty in Economics and Other Reflections*. Cambridge: Cambridge University Press, 1955.

Von Neumann, John and Morgenstern, Oskar. *Theory of Games and Economic Behavior*. Princeton: Princeton University Press, 1944.

Wald, Abraham. *Statistical Decision Functions*. New York: John Wiley and Sons, 1950.