

Università degli Studi di Firenze
Dipartimento di Psicologia

Dottorato di Ricerca in
Psicologia della Salute e dei Processi di Sviluppo
XVIII ciclo

La misurazione del rischio:
un contributo dalla
Teoria della Detezione del Segnale

Dottoranda

Dott.ssa Emilia Salvadori



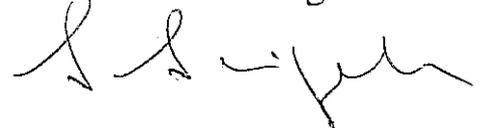
Tutor

Prof. Riccardo Luccio



Coordinatore

Prof. Saulo Sirigatti



Dicembre 2005

Indice

CAPITOLO 1	1
IL RISCHIO: DEFINIZIONE E MISURAZIONE	1
1.1 IL RISCHIO	1
1.2 LO STUDIO E LA MISURAZIONE DEL RISCHIO	5
1.2.1 APPROCCIO SPERIMENTALE	6
1.2.1.1 Teorie dell'utilità attesa	7
1.2.1.2 Teoria della razionalità limitata	11
1.2.2 APPROCCIO PSICOMETRICO	11
1.2.2 APPROCCIO QUALITATIVO	12
1.3 IL RISCHIO NELLA PROFESSIONE MEDICA	13
1.4 IL RISCHIO IN ADOLESCENZA	21
CAPITOLO 2	30
LA TEORIA DELLA DETEZIONE DEL SEGNALE	30
2.1 INTRODUZIONE	30
2.2 LA STORIA DELLA TDS	31
2.3 LA TEORIA DELLA DETEZIONE DEL SEGNALE E L'ANALISI ROC	36
2.3.1 MISURE DI ACCURATEZZA	43
2.3.2 MISURE DEL CRITERIO DI DECISIONE	50
2.3.3 CRITERIO OTTIMALE DI DECISIONE	54
CAPITOLO 3	61
LA RICERCA	61
3.1 INTRODUZIONE	61
3.2 PRESA DI DECISIONE IN AMBITO MEDICO	63
3.2.1 STUDIO 1	66
3.2.2 STUDIO 2	71
3.3 PRESA DI DECISIONE IN ADOLESCENZA	74
3.3.1 STUDIO 1	76
3.3.2 STUDIO 2	84
CAPITOLO 4	93
CONCLUSIONI	93
4.1 PRESA DI DECISIONE IN AMBITO MEDICO	93
4.2 PRESA DI DECISIONE IN ADOLESCENZA	94
4.3 CONCLUSIONI GENERALI	97
RIFERIMENTI BIBLIOGRAFICI	98

Capitolo 1

Il rischio: definizione e misurazione

1.1 Il rischio

Prendere una decisione significa scegliere tra una serie di alternative possibili rischiando di commettere un errore. Ogni azione o scelta può essere caratterizzata nei termini del grado di rischio ad essa associato. In senso molto generale, una azione è definita rischiosa quando vi è una certa probabilità che possa produrre conseguenze indesiderate. Tuttavia le definizioni del rischio sono molteplici e anche piuttosto diverse tra loro.

Alcune definizioni del rischio (**Approccio sperimentale**) si basano sulla teoria della probabilità ed hanno lo scopo di fornire una misura obiettiva del rischio, ad esempio nei termini di prodotto tra probabilità e grandezza della perdita (Brun, 1994). Questi approcci si diversificano per una serie di dettagli rilevanti: cosa si intende esattamente per probabilità e gravità, i metodi di misurazione delle stesse e il modello adottato per comporre queste componenti insieme (ad esempio moltiplicativo o additivo).

Altre definizioni del rischio (**Approccio psicometrico**) sono focalizzate sulla percezione soggettiva dello stesso ed hanno lo scopo di definire che cosa è il rischio per le persone comuni (Brun, 1994). Questo approccio più descrittivo parte dal concetto di probabilità come soggettivamente percepita. Yates e Stone (1992) considerano il rischio come un costrutto soggettivo che ha a che fare con: la perdita, il significato della perdita e l'incertezza ad essa associata. Anche in questo approccio il rischio è comunque considerato come l'effetto comune di probabilità della perdita e gravità della stessa, la differenza è che viene studiata la percezione soggettiva del rischio e quindi la probabilità diviene la probabilità soggettiva e la gravità è anch'essa determinata soggettivamente.

La probabilità è quindi una componente comune ad entrambi gli approcci al rischio e le differenze in termini di definizione della probabilità sono molto

importanti. Il problema potrebbe essere ridefinito nei termini dello scontro tra “frequentisti” e “soggettivisti”, e viene ampiamente discusso sia dalla letteratura sulla stima e percezione del rischio, sia dai teorici della probabilità (Brun, 1994).

Data la complessità dell’oggetto di studio, un problema importante è se sia veramente possibile formulare una definizione del rischio valida per tutte le situazioni. In altri termini è necessario rispondere alla domanda: che cosa intendono le persone comuni quando definiscono una situazione o una azione come rischiosa? In realtà, se si chiede alle persone comuni di definire il rischio, si ottengono risposte molto diverse, che dipendono fortemente dal contesto e dalla modalità con cui la questione è posta. Se il contesto è quello ad esempio della lotteria, o del gioco, il rischio riguarda l’assenza di certezza (Lopes, 1987), quando si parla di vaccinazioni o tecnologie le persone definiscono il rischio come la possibilità di riportare dei danni fisici (Slovic, 2000), infine quando si chiede ad adolescenti di riflettere su attività rischiose come fumare, bere alcolici o praticare attività sessuali, il rischio diventa il danno fisico e la possibilità di finire nei guai (Moore & Gullone, 1996). Il significato attribuito al termine rischio è talmente legato alla natura del particolare evento in questione, che la sua definizione dipende fortemente dal contesto in cui è utilizzata (Brun, 1994).

Un ulteriore quesito è se sia possibile fornire una definizione unica del rischio che sia adatta a tutti gli individui. Gli studi hanno fornito risultati controversi da questo punto di vista. Drottz-Sjoberg (1991), avendo fornito ai partecipanti quattro definizioni di rischio, trovò che il 45.8% rispondeva che il rischio è primariamente una questione di probabilità, mentre Brun (1994), avendo lasciato i suoi partecipanti liberi di fornire una definizione propria, riscontrò che il 62.5% lo definiva in termini di combinazione tra probabilità ed aspetti legati alle conseguenze. Un ulteriore risultato di questo secondo studio è il fatto che i partecipanti hanno fornito una percezione particolarmente individualistica del rischio, confermando peraltro risultati analoghi emersi in precedenza (Teigen, Brun & Slovic, 1988).

Un punto certo è che, per quanto rimanga aperta la questione di come misurare e combinare i due aspetti, la probabilità e la gravità delle conseguenze sono fattori fondamentali nella determinazione del rischio.

Un'ulteriore questione è che la stima del rischio data dalle persone comuni è sempre diversa da quella fornita dagli esperti ed è quindi presumibilmente influenzata da fattori ulteriori e specifici. Slovic (2000) ha ampiamente studiato la percezione del rischio nelle persone comuni rispetto ad una serie di attività (ad esempio la caccia, lo sci, il consumo di alcool, la guida, il nuoto, la scalata in montagna, ecc.) e tecnologie (ad esempio l'energia nucleare, i pesticidi, l'energia elettrica, i raggi x, i coloranti nei cibi, le vaccinazioni, ecc.), ed ha trovato che, mentre la stima degli esperti è largamente basata sulla frequenza di decessi causati dall'evento target, la stessa fornita dalle persone comuni risente di altre considerazioni. Slovic (2000) ha controllato la capacità delle persone comuni di fornire delle stime veritiere di frequenza di decessi ed i risultati mostrano che le stime sono buone. Di conseguenza, la differenza rispetto agli esperti non è data da una incapacità di stima oggettiva, bensì dall'intervento di altri fattori soggettivi nel giudizio, che vengono classificati e riassunti qui di seguito.

- Caratteristiche delle cause o origini del rischio. Tra queste è importante la distinzione tra cause naturali o umane, e tra cause indotte o accidentali. Laddove sia rintracciabile una responsabilità associata all'evento, esso viene percepito come maggiormente pericoloso (Brun, 1994).
- Caratteristiche dell'evento. Una caratteristica importante è l'universalità della fonte di rischio, in altri termini: quante persone sono a rischio. Da un lato i risultati mostrano che maggiore è il numero di persone esposte, maggiore è il rischio percepito (Slovic, Fischhoff & Lichtenstein, 1980), dall'altro è anche stato dimostrato che la percezione del rischio nelle persone comuni è connotata in senso individualistico (Teigen, Brun &

Slovic, 1988). Un'altra caratteristica associata all'evento è la novità dello stesso. Situazioni nuove e poco conosciute sono percepite come maggiormente rischiose e, se provocano un incidente, spesso si produce il cosiddetto effetto "punta dell'iceberg", cioè una conferma della sospetta pericolosità e la credenza che possano avere luogo danni anche più gravi (Brun, 1994).

- Caratteristiche delle conseguenze. Le caratteristiche relative alle conseguenze sono molteplici. Ad esempio il fatto che in realtà per talune azioni vi siano dei benefici. Nella vita di tutti i giorni vengono adottati comportamenti potenzialmente molto pericolosi (ad esempio guidare), che però comportano prima di tutto dei benefici. Inoltre le conseguenze hanno una distribuzione spaziale e temporale. Conseguenze negative che avvengono a persone vicine sono percepite peggiori che se avvenissero a sconosciuti, e conseguenze a breve termine sono anch'esse valutate come più rischiose che conseguenze a lungo termine (Brun, 1994). Una ulteriore caratteristica delle conseguenze è il disastro potenziale (Slovic, 2000). Nel caso di tecnologie particolari, come ad esempio l'energia nucleare, gli individui valutano con attenzione l'estensione del potenziale disastro. In generale, un evento catastrofico è percepito come peggiore rispetto a tanti piccoli incidenti. Peraltro in realtà non è fondamentale di per sé il numero di morti, ma viene anche valutato chi e come muore. Una modalità di morte molto sofferta o spaventosa è percepita come peggiore, soprattutto se sono coinvolti gruppi di soggetti vulnerabili e innocenti, ad esempio i bambini (Brun, 1994). Emozioni di forte paura collegate a conseguenze negative sono un fattore molto importante nella determinazione della rischiosità di un evento (Slovic, 2000).
- Caratteristiche della relazione della persona con l'evento. Il controllo degli eventi è un fattore certamente fondamentale. Le situazioni

percepite come incontrollabili sono anche giudicate come più pericolose: colui che si trova al volante percepisce come inferiore il rischio di un incidente rispetto ai passeggeri (Bragg & Finn, 1982). Un'altra caratteristica importante è la volontarietà dell'esposizione al rischio (Brun, 1994): le attività volontarie sono meglio accettate, anche se hanno lo stesso livello di pericolosità di quelle non volontarie. Infine determinante è anche la conoscenza del rischio (Johnson, 1993), per quanto la sua influenza sembra differire nelle diverse culture.

I diversi approcci nella definizione del rischio danno origine ad una serie di metodi diversi per la misurazione e lo studio dello stesso.

1.2 Lo studio e la misurazione del rischio

La ricerca nell'ambito della percezione del rischio si è concentrata su due questioni fondamentali (Brun, 1994):

- Quali sono i problemi che la gente comune considera più importanti, e se questi problemi sono diversi da quelli giudicati pericolosi dagli esperti e dalle autorità.
- I meccanismi psicologici in grado di spiegare le reazioni (atteggiamenti o azioni) della gente comune di fronte ai pericoli.

Esistono inoltre numerosi tipi di studi ciascuno con uno specifico paradigma di ricerca: alcuni indagano le reazioni delle persone comuni di fronte ad un determinato fenomeno; altri singole e specifiche attività rischiose; taluni studiano il rischio in senso ipotetico; alcuni chiedono ai partecipanti una stima personale, altri una stima del rischio collettivo, e così via (Brun, 1994).

Stabilire che cosa si intenda indagare non è peraltro una questione marginale. Il comportamento a rischio e la percezione del rischio sono collegati, ma non sono intercambiabili: la seconda è una delle determinanti del primo, ma non lo esaurisce. In senso generale potremmo dire che l'approccio sperimentale

si occupa di studiare il comportamento a rischio, mentre l'approccio psicometrico ha come oggetto la percezione del rischio.

Prima di entrare nell'ambito di questi specifici approcci, è necessario avere consapevolezza che la misurazione del rischio comporta, per sua natura, una serie di limitazioni: il modo in cui sono formulate le domande, la modalità di risposta, l'argomento ed il contesto della rilevazione, il fatto di concentrare l'attenzione su una sola situazione alla volta o di proporre più situazioni rischiose contemporaneamente, sono tutte variabili che influenzano pericolosamente l'attendibilità dei dati ottenuti (Johnson & Tversky, 1984; Slovic, 2000; Tversky & Koehler, 1994; Tyszka & Goszczyńska, 1993).

Nel descrivere i metodi di misurazione del rischio, oltre ai già menzionati approcci sperimentale e psicometrico, verrà brevemente introdotto anche un approccio di tipo qualitativo al problema.

1.2.1 Approccio sperimentale

Il metodo tradizionale di studiare il rischio nella presa di decisione è stato, in psicologia come in economia, quello della scommessa (Brun, 1994). In questo tipo di studi i partecipanti vengono sottoposti a prove del tipo lotteria, descritte in termini di vincite o perdite e di incertezza del risultato. A ciascuno viene chiesto di scegliere tra le varie lotterie o di valutarne la rischiosità. Questi approcci hanno il vantaggio di essere semplici e, al tempo stesso, robusti dal punto di vista metodologico.

I primi teorici della scelta si sono basati sulla Teoria dell'utilità attesa: l'ipotesi di partenza è che l'individuo tenta di massimizzare i benefici, ottimizzando quindi i risultati. In seguito i teorici si sono orientati verso un modello parzialmente diverso, quello della razionalità limitata: l'assunto di base è che gli esseri umani, piuttosto che ottimizzare a tutti i costi, si accontentano di raggiungere un soddisfacimento, ovvero un compromesso accettabile tra assicurarsi degli esiti positivi ed evitare delle perdite eccessive (Brun, 1994).

I risultati degli studi basati sull'approccio sperimentale sono stati considerati generalizzabili al mondo reale ed utilizzati come punto di partenza per costruire modelli del comportamento a rischio (Lopes, 1987).

Ciò che è stato criticato a questo approccio è che la sua semplicità è stata confusa con la generalizzabilità dei risultati trovati: nella vita reale le persone non si trovano davanti opzioni di scelta così ben definite, le vincite e le perdite sono multidimensionali (fisiche, psicologiche, economiche, ecc.) e le scelte spesso avvengono con poco tempo a disposizione (Brun, 1994).

1.2.1.1 Teorie dell'utilità attesa

La **Teoria classica dell'utilità attesa** (*Expected Utility Theory, EU Theory*) è stata la prima ad essere sviluppata per spiegare le scelte rischiose (Edwards, 1992).

Le decisioni prese in condizione di incertezza vengono tipicamente rappresentate da una tabella di costi (Tabella 1.1), in cui le righe corrispondono alle azioni alternative che il decisore può intraprendere e le colonne corrispondono ai possibili stati del mondo (aventi ciascuno una probabilità di occorrenza).

Tabella 1.1
Tabella dei costi

<u>Decisione</u>	<u>Stato del mondo</u>	
	Sole [p=.6]	Pioggia [p=.4]
Portare l'ombrello	Essere asciutto con l'ombrello (+1)	Essere asciutto con l'ombrello (+1)
Non portare l'ombrello	Essere asciutto senza l'ombrello (+2)	Essere bagnato senza l'ombrello (0)

Le celle della matrice contengono: le possibili conseguenze derivanti dall'occorrenza congiunta della decisione e dello stato del mondo; e l'utilità o desiderabilità soggettiva delle stesse.

Nel 1738, Bernoulli ha definito la scommessa migliore come quella che massimizza l'utilità attesa della decisione:

$$EU(A) = \sum_{i=1}^n P(E_i)U(X_i) \quad (1.1)$$

dove $EU(A)$ rappresenta l'utilità attesa di una azione che ha come conseguenze X_1, X_2, \dots, X_n che dipendono dagli eventi E_1, E_2, \dots, E_n , $P(E_i)$ rappresenta la probabilità che l'esito i -esimo di quella azione e $U(X_i)$ rappresenta il valore soggettivo o l'utilità di quella conseguenza.

Utilizzando i dati contenuti nella Tabella 1.1, è possibile, ad esempio, ricavare le utilità attese delle due possibili decisioni: prendere o meno l'ombrello. Nel calcolo relativo all'utilità attesa della decisione di portare l'ombrello, in base alla formula 1.1, dovrà essere effettuata la sommatoria tra: la probabilità dell'evento sole (.6) moltiplicata per l'utilità della conseguenza [sole + portare l'ombrello] (+1) e la probabilità dell'evento pioggia (.4) moltiplicata per l'utilità della conseguenza [pioggia + portare l'ombrello] (+1):

$$EU(\text{ombrello}) = [.6(+1)] + [.4(+1)] = 1.0 \quad (1.2)$$

Svolgendo la formula 1.2, risulta che l'utilità attesa della decisione di portare l'ombrello è pari a 1.

Per calcolare l'utilità attesa della decisione di non portare l'ombrello si procede analogamente sommando: probabilità dell'evento sole (.6) moltiplicata per l'utilità della conseguenza [sole + non portare l'ombrello] (+2) e probabilità dell'evento pioggia (.4) moltiplicata per l'utilità della conseguenza [pioggia + non portare l'ombrello] (0):

$$EU(\text{no - ombrello}) = [.6(+2)] + [.4(0)] = 1.2 \quad (1.3)$$

Il risultato è che l'utilità attesa relativa alla decisione di non portare l'ombrello è di 1.2, cioè maggiore rispetto a quella di portarlo con sé. La decisione finale dovrebbe quindi essere che lasciare a casa l'ombrello è più vantaggioso che portarlo.

Un contributo importante alla teoria della decisione deriva da Von Neumann e Morgenstern (1947) che hanno messo a punto una giustificazione formale del criterio di utilità attesa: secondo questi studiosi, se le preferenze di un individuo soddisfano alcuni assiomi di base del comportamento razionale, allora le sue decisioni possono essere descritte in termini di massimizzazione dell'utilità attesa. In seguito Savage (1954) ha generalizzato la teoria, arrivando a definire $P(E_i)$ come il valore di probabilità soggettiva stabilita dal decisore.

I principi assiomatici da cui parte la teoria della massimizzazione dell'utilità attesa sono estremamente razionali. Uno di questi, la transitività, prevede che se il decisore preferisce l'alternativa A alla B e la B alla C, allora dovrà preferire anche l'alternativa A alla C; un altro stabilisce che gli esiti che non sono influenzati dalla decisione non dovrebbero, a loro volta, determinare la stessa. Questa teoria assume, quindi, che un decisore razionale preferisce selezionare quella opzione che è logicamente coerente sia con le sue preferenze per quanto riguarda gli esiti, che con la sua percezione della probabilità degli eventi.

Il **modello di Valori Attesi** (*Expected Values, EV*) di Atkinson (1957) si propone come uno sviluppo dei modelli di utilità attesa, che tiene conto anche delle differenze individuali. Secondo questo autore, ogni individuo presenta due motivazioni in competizione: una motivazione al successo (M_s) e una motivazione all'evitamento del fallimento (M_{ef}).

Oltre alle motivazioni, il modello include altre quattro variabili: la probabilità soggettiva o aspettativa di successo (P_s), la probabilità soggettiva di fallimento (P_f), il valore incentivo del successo (I_s), il valore incentivo

(negativo) del fallimento (I_f). Le variabili sono combinate per ottenere una misura della motivazione risultante:

$$\text{Motivazione} = (M_s \times P_s \times I_s) + (M_{ef} \times P_f \times -I_f) \quad (1.4)$$

Il modello prevede che le differenze individuali producano in ciascuno una diversa forza, e quindi un diverso bilanciamento, di queste due motivazioni. Individui orientati al successo ($M_s > M_{ef}$) tenderanno a preferire opzioni moderatamente rischiose, mentre individui orientati all'evitamento del fallimento ($M_s < M_{ef}$) preferiranno opzioni poco rischiose (dove il successo è quasi certo) o molto rischiose (dove il fallimento può essere spiegato dall'estrema difficoltà della situazione).

Questo modello ha ottenuto degli esiti abbastanza soddisfacenti in studi che riguardavano compiti di abilità orientati alla vittoria, mentre i risultati sono stati piuttosto inconsistenti laddove la situazione sperimentale richiedeva una scelta basata sulla probabilità (Slovic, 1964).

Uno sviluppo successivo derivato dalla Teoria classica dell'utilità attesa, è la **Teoria Prospettica** (*Prospect Theory*) sviluppata da Kahneman e Tversky nel 1979. Gli obiettivi fondamentali di questo approccio sono tre:

- Spiegare perchè la maggior parte delle persone preferiscono scelte sicure, anzichè scelte rischiose aventi lo stesso valore atteso di utilità (*certainty effect*).
- Spiegare perchè la maggior parte delle persone manifestano avversione al rischio nel caso di vincita, ma propensione al rischio nel caso di perdita (*reflection effect*).
- Spiegare perchè la maggior parte delle persone si focalizzano sulle differenze tra le opzioni di scelta e ignorano gli aspetti comuni (*isolation effect*).

1.2.1.2 Teoria della razionalità limitata

Simon (1956, 1959) propone questa teoria in contrapposizione all'approccio di massimizzazione dell'utilità attesa. Lo studioso rileva che quest'ultimo prevede che l'uomo si trovi a scegliere tra un numero fisso e determinato di alternative, ciascuna avente delle conseguenze anch'esse esplicite ed evidenti. Questo nel mondo reale non avviene, data tutta una serie di fattori intervenienti, interni ed esterni al decisore, che comunque ne influenzano l'agire.

La teoria della razionalità limitata parte dal presupposto che l'individuo non pensi in modo probabilistico, ma che altresì cerchi di evitare la necessità di affrontare direttamente l'incertezza. Quindi le limitazioni cognitive del decisore lo costringono, per poter gestire le varie situazioni, a costruire un modello semplificato del mondo e ad agire in base a strategie del tipo per prove ed errori. Il principio di base di questa teoria è quello di soddisfacimento: l'individuo si accontenta di ottenere un livello di controllo soddisfacente, e non necessariamente ottimale (Slovic, 2000).

1.2.2 Approccio psicometrico

Questo approccio, fondato da Slovic, Lichtenstein e Fischhoff (1984), parte dall'assunto che il rischio sia, per sua stessa natura, soggettivo, che non esista in quanto tale, ma che sia un concetto creato dalla mente umana per comprendere e gestire i pericoli e le incertezze della vita (Brun, 1994). Inoltre il rischio viene definito come complesso e multidimensionale (Slovic, 2000) e la sua misurazione è focalizzata sulle preferenze espresse dagli individui.

Il punto di partenza è stato quello di voler comprendere le dimensioni o caratteristiche del rischio per come è percepito dalle persone comuni (Slovic, 2000). Dopo aver, mediante una revisione della letteratura, creato una lista di eventi rischiosi e di dimensioni correlate alla valutazione del rischio, queste ultime sono state verificate empiricamente mediante una serie di studi. Le

dimensioni del rischio sono state sottoposte ad analisi fattoriali esplorative, per ridurle ad un numero inferiore di fattori sovraordinati. Il pattern tipico, e replicato peraltro abbastanza stabilmente da quasi tutti gli studi, include due fattori: rischio spaventoso (*dread risk*) e rischio sconosciuto (*unknown risk*) (Slovic, 2000). La procedura prevede poi che venga costruito uno spazio fattoriale a partire da questi due fattori, e che, al suo interno, vengano collocati i vari eventi analizzati, in modo da poterne confrontare la posizione rispetto ai due assi (Slovic, 2000).

Per quanto riguarda la applicabilità di questo approccio, gli studi dimostrano che si riscontra una buona stabilità a livello interindividuale, culturale e di tipo di fenomeno indagato per quanto concerne la rappresentazione bifattoriale del rischio, mentre, se si considera il livello di rischiosità associato ai vari eventi, questa stabilità decade per culture diverse, gruppi sociali e genere (Brun, 1994; Slovic, 2000).

Le critiche mosse a questo approccio sono sostanzialmente due: la natura correlazionale della metodologia applicata e la mancanza di rilevazione della variabilità tra individui.

1.2.2 Approccio qualitativo

Gli studiosi che hanno provato a studiare il rischio mediante approcci più qualitativi, hanno solitamente richiesto ai partecipanti di stilare una lista di rischi considerati rilevanti (Brun, 1994).

Il risultato più interessante è che, laddove l'approccio psicometrico aveva individuato una maggiore paura per le nuove tecnologie rispetto ai rischi di tutti i giorni, questo risultato si è sostanzialmente ribaltato quando ai partecipanti è stato chiesto di comporre una loro propria lista. In questo senso, Fischer, Morgan, Fischhoff, Nair e Lave (1991) hanno riscontrato che la categoria incidenti è la più menzionata (37%) e le malattie la seconda; ed analogamente

Brun (1994) ha trovato che malattie ed incidenti comprendevano complessivamente il 68% delle risposte.

Come è risultato evidente dai paragrafi precedenti, il rischio può essere studiato da prospettive anche molto diverse tra loro e quindi utilizzando metodi disparati. Molti studi psicologici hanno adottato una prospettiva individualistica, cercando di spiegare la percezione del rischio ed il comportamento a rischio in termini di strategie o abilità cognitive, tratti di personalità ed esigenze situazionali. Questo approccio allo studio del rischio dipende fortemente dall'oggetto di studio e dalle specificità degli individui e dei fenomeni indagati. In questo senso esso verrà considerato durante la trattazione degli ambiti specifici che questa ricerca ha voluto studiare: il rischio nella professione medica ed in adolescenza.

1.3 Il rischio nella professione medica

Il medico è tenuto per lavoro a compiere scelte importanti e, data la mancanza di tutte le informazioni necessarie e spesso del tempo sufficiente a decidere, opera nell'ambito della decisione presa in condizioni di incertezza. La medicina rappresenta infatti un ambito ideale per lo studio della presa di decisione: la scelta è solitamente tra un ristretto numero di opzioni, vi è un certo grado di incertezza riguardo agli esiti e vi sono molti parametri rispetto ai quali confrontare l'efficacia delle opzioni di risposta (Chapman & Sonnenberg, 2000).

Nella presente ricerca, l'interesse è rivolto specificamente alla decisione del medico rispetto all'impatto che la diagnosi ha sul paziente, di conseguenza non verranno presi in considerazione alcuni aspetti molto studiati in ambito medico: le decisioni del paziente o le decisioni del medico rispetto ai costi sanitari o all'efficacia di taluni test diagnostici o trattamenti.

Nello studio della presa di decisione in ambito medico vengono solitamente utilizzati modelli che si basano sulla teoria classica della decisione: nello

specifico è stata ampiamente utilizzata la teoria dell'utilità attesa (*Expected Utility Theory*), specifica per la scelta in condizioni di incertezza (Chapman & Sonnenberg, 2000).

I metodi quantitativi, elaborati in ambito medico a partire da queste teorie, partono dall'assunto che la presa di decisione sia da considerarsi un processo nel quale intervengono variabili quali la probabilità degli esiti ed il valore degli stessi (Chapman & Sonnenberg, 2000).

Per costruire i modelli di misurazione, e quindi curare l'aspetto quantitativo della ricerca, è però fondamentale rispondere preliminarmente ad una questione più prettamente qualitativa: come vengono prese le decisioni in ambito medico? La ricerca in questo ambito si è interessata a due fattori che possono influenzare questo processo: le strategie cognitive che il medico utilizza per raccogliere e sintetizzare tutte le informazioni che lo conducono alla decisione finale, alcune caratteristiche specifiche della professione medica, quali ad esempio l'esperienza lavorativa.

Un modo per esplorare le strategie cognitive utilizzate dal medico, è quello di confrontare l'accuratezza della decisione con quanto sarebbe previsto a partire dalle norme della teoria della probabilità (Chapman & Sonnenberg, 2000), in altre parole: arrivare a comprendere gli elementi che determinano la presa di decisione del medico partendo dagli errori e dalle violazioni che esso compie rispetto ad un corretto ragionamento probabilistico.

Le deviazioni dagli standard normativi sono importanti per due ragioni: offrono esempi dei processi cognitivi che sottostanno alla presa di decisione e indicano quelle aree dove è necessario un miglioramento (Chapman & Elstein, 2000). Dal punto di vista cognitivo, esistono tre possibili fonti di errore nella decisione del medico: 1) giudizio di probabilità di particolari eventi, quali l'occorrenza di una diagnosi o degli effetti di un trattamento; 2) valutazione dell'utilità, preferenza o scelta di un esito; 3) tempo a disposizione e sequenza temporale degli eventi.

Il problema relativo al giudizio di probabilità degli eventi è connesso al fatto che molte decisioni avvengono in condizioni di incertezza. Spesso i medici devono compiere delle stime intuitive della probabilità di una diagnosi o dell'efficacia di una terapia basandosi sulle informazioni disponibili e possono incorrere in alcuni errori.

Un errore molto comune è quello che la Teoria del Supporto (Tversky & Koehler, 1994) definisce come principio del disfaccimento: fornire una descrizione più dettagliata di un evento aumenta la sua probabilità stimata. Redelmeier, Koehler, Liberman e Tversky (1995) hanno infatti dimostrato che rappresentazioni alternative dello stesso evento possono dar luogo a diversi giudizi di probabilità. In particolare, la stima di probabilità di una diagnosi varia in base alla lunghezza della lista di possibili patologie connesse allo stesso complesso di sintomi. Nella condizione in cui vengono offerte cinque possibili diagnosi (*long list condition*), la probabilità viene distribuita tra le alternative, mentre se vengono fornite solo due possibili diagnosi (*short list condition*) avviene una sovrastima della probabilità (Redelmeier et al, 1995).

L'errore retrospettivo (*hindsight bias*), interviene quando il decisore gonfia la probabilità di una diagnosi corretta (Fischhoff, 1975). Arkes, Wortmann, Saville e Harkness (1981) hanno presentato a cinque gruppi di medici un caso clinico con quattro potenziali diagnosi. Ai medici del gruppo di controllo è stato chiesto di stimare la probabilità di ciascuna delle quattro diagnosi, ai medici dei restanti gruppi è stata comunicata la diagnosi corretta (una diversa per ciascun gruppo) e dopo è stato chiesto loro di stimare la probabilità che avrebbero assegnato alle quattro possibili diagnosi se non avessero saputo quella giusta. I medici dei quattro gruppi sperimentali hanno distorto le loro stime in base alla diagnosi che era stata loro presentata come corretta. Molti studi hanno inoltre dimostrato che la conoscenza dell'esito focalizza l'attenzione del decisore sulle informazioni che sono coerenti con lo stesso e rendono quella diagnosi sempre più probabile, portandolo ad ignorare invece tutte le informazioni che

renderebbero plausibili diagnosi alternative (Chapmann & Elstein, 2000). Evidenze coerenti con la diagnosi sono ricordate più facilmente rispetto a fatti che contraddicono l'esito (Dellarosa & Bourne, 1984).

Il terzo possibile errore relativo al giudizio di probabilità degli eventi è l'errore confermativo (*confirmation bias*) (Chapmann & Elstein, 2000). Un aspetto importante del ragionamento diagnostico è la verifica di ipotesi: i medici devono determinare quale, tra molte diagnosi (o ipotesi), è più probabile data l'evidenza ottenibile dalla storia del paziente, gli esami ed i test di laboratorio. La teoria normativa per la verifica di ipotesi è il Teorema Bayesiano, ma nel processo di decisione gli individui spesso non rispettano questa legge e possono compiere numerosi errori: ad esempio sovrastimare un'ipotesi particolare, poiché fuorviati da una credenza pregressa sulla sua veridicità; oppure interpretare i dati in modo errato, sottovalutando le informazioni a disconferma della stessa.

Passando agli errori connessi alla valutazione dell'utilità di una opzione di scelta, il problema fondamentale è che le preferenze non sono stabili, bensì sono influenzate dal modo in cui vengono elicitate (Slovic, 1995).

L'esempio più noto della labilità delle preferenze è l'effetto del contesto (*framing effect*) (Kahnemann & Tversky, 1984). Questo effetto viola il principio normativo della invarianza descrittiva, secondo il quale descrizioni equivalenti della stessa situazione dovrebbero portare alla stessa scelta. A causa dell'effetto del contesto, il decisore valuta gli esiti in termini di guadagno o di perdita in base alla modalità con cui la situazione viene proposta: se, ad esempio, uno scenario è descritto nei termini di numero di vite salvate si preferisce ridurre il rischio e massimizzare il guadagno; mentre se lo stesso identico scenario è descritto nei termini di vite perse, si preferisce compiere scelte rischiose ed azzardare per ridurre le perdite. L'effetto del contesto è stato indagato in molti studi sulla presa di decisione in ambito medico (Chapmann & Elstein, 2000), ma i risultati non sono univoci e non è ancora chiaro quali caratteristiche della decisione determinano questo effetto.

Un secondo errore connesso alla scelta è l'inversione della preferenza (*preference reversals*) (Chapmann & Elstein, 2000). Esso viola il principio normativo della invarianza procedurale: metodi diversi di elicitare un preferenza dovrebbero fornire lo stesso ordine di preferenza. In altri termini: se una persona sceglie A al posto di B, dovrebbe essere disponibile a pagare di più per A che per B, e dovrebbe dare una valutazione migliore di A che di B. Molti studi hanno tuttavia dimostrato che spesso questo principio viene violato (Chapmann & Elstein, 2000; Slovic & Lichtenstein, 1983) e che le procedure e le domande usate per elicitare un preferenza sono fattori che influenzano questo fenomeno.

Un terzo errore connesso all'espressione di preferenze è l'aggiunta di alternative nella decisione. Un medico può trovarsi a decidere tra due trattamenti A e B, supponiamo che una terza alternativa di trattamento (C) venga aggiunta alle possibilità di risposta. Secondo il principio normativo della regolarità, l'aggiunta di C non dovrebbe aumentare la probabilità nella scelta di A o B, dal momento che l'utilità attesa di ogni opzione non è influenzata dal numero di opzioni disponibili, ma questo principio viene spesso violato (Chapmann & Elstein, 2000). Secondo Shafir, Simonson e Tversky (1997), la scelta può essere compresa sulla base delle ragioni che l'individuo usa per spiegare e giustificare la decisione.

L'errore di omissione (*omission bias*) deriva da una violazione della consequenzialità della teoria delle decisioni: la decisione ottimale è quella che, al di là di tutto, ottiene le migliori conseguenze, mentre il tipo di azione effettuata per ottenere le conseguenze è irrilevante (Chapmann & Elstein, 2000). Nel caso di una patologia anche grave, se la strategia di attesa ottiene dei risultati migliori rispetto ad un'operazione chirurgica, allora è la decisione ottimale. Il fatto che si tratti di una strategia passiva dovrebbe essere irrilevante ai fini della decisione. Eppure l'evidenza che l'operazione chirurgica rimane la procedura più frequentemente suggerita ai pazienti, lascia supporre che la valutazione dei medici possa essere influenzata da altri fattori: probabilmente

come professionisti preferiscono le azioni piuttosto che le omissioni (Chapmann & Elstein, 2000). In altri casi, tuttavia (Baron & Ritov, 1994; Chapmann & Elstein, 2000), si è riscontrato che il decisore considera le omissioni che portano a conseguenze negative come meno immorali e meno negative, rispetto ad azioni che portano alle medesime conseguenze.

L'incertezza non è l'unico fattore cognitivo che induce la decisione medica all'errore, anche il tempo causa una serie di *bias*.

La teoria dell'utilità scontata è la teoria normativa delle scelte tra esiti che avvengono in momenti diversi nel futuro. Il valore assoluto degli esiti futuri deve essere corretto sulla base di un valore noto come rata di sconto, che dovrebbe essere costante e positiva. In realtà le rate di sconto utilizzate implicitamente per le decisioni sanitarie sono molto instabili: spesso molto alte, in altri casi pari a 0 o addirittura negative (Chapmann & Elstein, 2000). Le rate di sconto soggettive sono poi influenzate da numerosi fattori non normativi che rappresentano un ulteriore problema: sono più alte per i guadagni che per le perdite, per esiti a breve termine piuttosto che a lungo termine (Chapmann & Elstein, 2000).

La teoria dell'utilità scontata serve anche per la valutazione delle sequenze di eventi: ogni esito nella sequenza deve essere corretto in base al tempo trascorso e il valore totale della sequenza risulta dalla somma dei valori corretti. Nella valutazione degli esiti individuali, il decisore generalmente preferisce avere un buon esito a breve termine, piuttosto che a lungo termine (preferenza positiva del tempo), mentre nella valutazione di sequenze, il decisore di solito preferisce incrementare la sequenza, piuttosto che avvicinare l'esito (preferenza negativa del tempo).

Data la quantità e qualità dei possibili errori nella presa di decisione in medicina, e la gravità dell'impatto di una decisione errata da parte del personale medico, sia sui costi del sistema sanitario, sia per il paziente, è evidente che

riuscire a determinare i fattori cognitivi che influenzano e determinano la presa di decisione del medico è un obiettivo importante nello studio della decisione.

Il secondo tipo di fattori che hanno una influenza molto forte sulla presa di decisione in medicina riguardano più da vicino questa specifica professione. L'esperienza lavorativa ed il feedback che il professionista riceve nel tempo rispetto alle proprie decisioni, sono determinanti importanti della percezione di accuratezza del proprio operato e della sensazione di fiducia nelle proprie decisioni.

L'esperienza, in ambito medico, può essere determinata da molte caratteristiche quali: età, livello di preparazione, numero di anni di pratica o precedenti esperienze con fenomeni specifici come, ad esempio, il numero di pazienti che sono deceduti. Berwick e Thibodeau (1983) hanno riscontrato che nei pediatri il livello di preparazione influisce sulla accuratezza di diagnosi per una certa patologia, ma non per altre. Dawson, Connors, Speroff, Kemka, Shaw e Arkes (1993) non hanno riscontrato, per quanto riguarda l'accuratezza diagnostica dello stato emodinamico, differenze in base al livello di esperienza. Infine, Dawson (2000) ha esaminato la relazione tra un serie di caratteristiche dei medici (molte delle quali relative all'esperienza lavorativa) e l'accuratezza nella stima di sopravvivenza di pazienti adulti, ospedalizzati e seriamente malati. I risultati mostrano che, di tutte le caratteristiche esaminate, soltanto in grado di esperienza con i pazienti deceduti è significativamente associato con l'accuratezza della stima in fase di prognosi.

I risultati di questi studi sembrano condurre a tre possibili conclusioni (Dawson, 2000): 1) il livello generale di esperienza del medico non sembra avere effetto sull'accuratezza del giudizio riguardo la prognosi; 2) il grado di esperienze specifiche (ad esempio l'esperienza precedente con pazienti deceduti) potrebbe essere un fattore importante per la valutazione prognostica; 3) al momento della diagnosi il livello generale di esperienza potrebbe giocare un ruolo importante nell'accuratezza del giudizio, ma questo effetto deve essere

ulteriormente studiato e verificato, in particolare la sua possibile dipendenza dal tipo specifico di malattia.

Non solo il livello di esperienza in generale, ma anche il feedback è un fattore importante, mediante il quale l'esperienza può portare ad una maggiore accuratezza del giudizio. Purtroppo, nella pratica medica di ogni giorno, entrano in gioco considerazioni etiche, economiche o pragmatiche che talvolta possono fornire un feedback inadeguato o fuorviante rispetto alla presenza di malattia (Dawson, 2000). Una sequenza non perfetta di test diagnostici e dei risultati parziali possono contribuire alla scarsa accuratezza nella diagnosi di alcune patologie.

Infine, per quanto l'accuratezza può o meno migliorare con l'esperienza, la confidenza, intesa come grado di fiducia nel proprio giudizio, spesso aumenta. Einhorn e Hogarth (1978) hanno suggerito che la confidenza è direttamente proporzionale al numero precedente di giudizi espressi, indipendentemente dalla correttezza degli stessi. Connors, Dawson, Speroff, Arkes, Knaus, Harrel, Lynn, Terro, Goldman, Califf, Fulkerson, Oye, Bellamy e Desbiens (1992) hanno riscontrato che l'accuratezza prognostica dei medici varia in modo proporzionale alla confidenza degli stessi. In modo contrastante, Dawson e colleghi (1993) hanno riscontrato che la confidenza nella stima dello stato emodinamico incrementa significativamente nei cinque anni di specializzazione, mentre il livello di accuratezza non ha un andamento analogo.

La confidenza con cui viene espresso un giudizio ha poi una influenza diretta sui comportamenti adottati dal professionista: alti livelli di confidenza portano ad agire sulla base delle proprie valutazioni, mentre un basso livello di confidenza induce il professionista alla ricerca di informazioni addizionali e costituisce una buona opportunità per indurlo ad utilizzare modelli statistici predittivi e validati (Dawson, 2000).

Concludendo, i risultati sperimentali mostrano che la relazione tra accuratezza e confidenza non è ancora del tutto chiarita. Allo stesso tempo

l'importanza che questo aspetto riveste nella presa di decisione del professionista è evidente: la fiducia nel proprio giudizio spesso determina la presa di decisione e risulta quindi necessario verificare se si tratta o meno di un indicatore affidabile.

1.4 Il rischio in adolescenza

L'adolescenza è uno stadio della vita piuttosto ad alto rischio (Jessor, 1984). E' una fase intermedia, di passaggio, durante la quale gli individui sperimentano molti comportamenti, alcuni dei quali potenzialmente lesivi della loro incolumità.

I comportamenti usualmente considerati a rischio in adolescenza e, per questo, ampiamente studiati sono: incidenti stradali, uso e abuso di alcool, uso e abuso di sostanze stupefacenti, tabagismo e comportamenti sessuali non protetti (che quindi possono comportare malattie sessualmente trasmissibili e gravidanza).

Un riscontro comune a molti studi è che il coinvolgimento degli adolescenti nei comportamenti a rischio incrementa con l'età (Greening & Stoppelbein, 2000; Hampson, Severson, Burns, Slovic & Fisher, 2001; Slovic, 2000).

Un aspetto che ha fornito, invece, risultati controversi in molti studi riguarda le differenze di genere nella percezione del rischio e nella messa in atto di comportamenti rischiosi. Alcuni studi hanno mostrato una differenza di genere nella valutazione e propensione al rischio: in particolare i maschi sembrano essere più portati ad adottare comportamenti rischiosi e meno consapevoli dei rischi rispetto alle femmine (Donohew, 1988; Greene, Krcmar, Walters, Rubin & Hale, 2000; Newcomb & McGee, 1989; Tyler & Lichtenstein, 1997). In modo alquanto diverso, Savadori e Rumiati (1996) hanno riscontrato che le differenze di genere nella stima del rischio relativo ad una serie di comportamenti sono limitate ad alcune azioni specifiche, considerate culturalmente come maschili e femminili. I maschi considerano meno pericolosi

comportamenti a loro noti, quali gli scontri nello stadio tra tifoserie e l'utilizzo di fuochi di artificio; mentre le femmine sono maggiormente preoccupate per azioni quali comportamenti alimentari scorretti, rapporti sessuali non protetti e autostop, fenomeni tipicamente associati all'universo femminile. Infine Hampson, Severson, Burns, Slovic e Fisher (2001) hanno costruito un modello di equazioni strutturali che contenesse una serie di variabili in grado di predire il comportamento di consumo di sostanze alcoliche e guida in stato di ebbrezza. Il modello definitivo è stato testato separatamente nei due sottocampioni di maschi e femmine per controllare l'invarianza dello stesso: il risultato mostra che non vi è una differenza di genere tra le relazioni che legano tra loro le variabili in grado di predire il comportamento.

La percezione del rischio da parte di coloro che adottano comportamenti potenzialmente dannosi (*participants*) è molto diversa da quella di coloro che non praticano questo tipo di comportamenti (*non-participants*) (Benthin, Slovic & Severson, 1993; Brown, Christiansen & Goldman, 1987; Savadori & Rumiati, 1996; Slovic, 2000). In uno studio molto interessante sull'assunzione di rischio in adolescenza, Benthin, Slovic e Severson (1993) hanno riscontrato un vero e proprio profilo che diversifica i *participants* dai *non-participants*. I primi riportano una maggiore conoscenza dei rischi, meno paura, una minore percezione del rischio per sé stessi e le conseguenze come meno dannose, di possedere un maggiore controllo del rischio, e di risentire maggiormente dell'influenza dei pari. Sulla stessa linea sono i risultati riportati da Savadori e Rumiati (1996): i *participants* riportano maggiore influenza dei pari, sottostima della probabilità di incorrere in problemi, condivisione del comportamento da parte dei coetanei e maggiori benefici associati al comportamento. Slovic (2000) ha studiato in modo approfondito la relazione tra la messa in atto di un comportamento a rischio e la qualità delle aspettative che la persona ha rispetto a quello stesso comportamento. Questo autore (Slovic, 2000) ha riscontrato che coloro che praticano comportamenti a rischio sono maggiormente portati ad

associare al comportamento stesso valori ed esiti positivi e poco propensi ad associarvi valori ed esiti negativi, rispetto ai *non-participants*.

Inoltre spesso gli adolescenti che intraprendono un comportamento a rischio sono più predisposti ad adottarne anche altri. Jessor e Jessor (1977) hanno individuato questo problema all'interno della Teoria del Comportamento Problematico (*Problem Behavior Theory*) e lo hanno definito sindrome del comportamento problematico (*syndrome of adolescent problem behavior*). Jessor (1984, 1990) ha rilevato che l'uso di droghe è positivamente correlato con l'uso di alcool, con il tabagismo e con i comportamenti sessuali a rischio. Molti altri studi (Arnett & Balle-Jensen, 1993; Benthin, Slovic & Severson, 1993; Savadori & Rumiati, 1996; Slovic, 2000) hanno fornito evidenze di un quadro complesso di correlazioni tra vari tipi di comportamenti a rischio in adolescenza.

Data l'importanza del problema dei comportamenti a rischio in adolescenza, che possono comportare problemi anche molto seri per la salute, le campagne di informazione sono state molto numerose e mirate ai canali comunicativi giovanili. Tuttavia, la maggiore informazione non sembra aver cambiato la propensione dei giovani al rischio. Questa constatazione porta con sé una riflessione più profonda sul perché gli adolescenti siano così portati al rischio, sul significato che per gli adolescenti hanno i comportamenti a rischio, sulla loro percezione dello stesso, sugli eventuali errori di valutazione e sui fattori che influenzano le loro decisioni.

I fattori cognitivi che sono stati chiamati in causa per spiegare l'assunzione di rischio in adolescenza sono molteplici.

Una caratteristica che discrimina i ragazzi che praticano comportamenti a rischio da coloro che non lo fanno è una sovrastima della frequenza dei comportamenti nei pari (Benthin, Slovic & Severson, 1993). Il fattore cognitivo che induce questo tipo di errore è probabilmente una applicazione errata dell'euristica della disponibilità (Tversky & Kahneman, 1973): il comportamento sembra più diffuso perché, praticandolo, è semplicemente più

saliente. L'importanza delle euristiche, e soprattutto di un loro errato utilizzo nella presa di decisione in situazioni di rischio, è ampiamente studiata e riconosciuta in letteratura (Slovic, 2000).

Una serie di studi discute l'influenza di fenomeni quali l'egocentrismo, il *personal fable* e l'ottimismo ingiustificato (*optimistic bias*) sull'assunzione di rischio in adolescenza. L'egocentrismo ed il *personal fable* sono forme di pensiero abbastanza tipiche in adolescenza, e portano i ragazzi a sentirsi invulnerabili, unici ed immortali (Benthin, Slovic & Severson, 1993; Elkind, 1967; Savadori & Rumiati, 1996). Di conseguenza si ritengono meno a rischio rispetto ai coetanei che adottano lo stesso tipo di comportamento, e mostrano un atteggiamento ottimistico ingiustificato ed errato che fa loro sottovalutare l'impatto reale dei rischi per la loro incolumità (Savadori & Rumiati, 1996; Slovic, 2000). L'influenza di questi fattori è stata riscontrata in studi sull'uso di alcool (Greene, Krcmar, Walters, Rubin & Hale, 2000; Hansen, Raynor & Wolkenstein, 1991), sulla guida in stato di ebbrezza (Arnett, 1990a; 1992), sull'uso di sostanze stupefacenti (Arnett, 1992; Greene, Krcmar, Walters, Rubin & Hale, 2000) e sui comportamenti sessuali (Elkind, 1967; Greene, Krcmar, Walters, Rubin & Hale, 2000; Kalmuss, 1986).

La teoria classica dello sviluppo cognitivo collega il vissuto del *personal fable* all'atteggiamento di egocentrismo che accompagna la transizione al pensiero formale (Elkind, 1967) e questo permette di introdurre un ulteriore fattore cognitivo che potrebbe influenzare la percezione del rischio in adolescenza: lo sviluppo del pensiero formale, caratterizzato dal ragionamento astratto e ipotetico deduttivo (Inhelder & Piaget, 1958). Alcuni studi hanno infatti mostrato che gli adolescenti, che non hanno ancora acquisito piena abilità nel ragionamento formale, sono meno in grado di pianificare o posticipare la gratificazione del bisogno (Damody, 1991).

Oltre ai fattori cognitivi, sono stati presi in considerazione fattori quali la personalità degli adolescenti, la motivazione e la sfera morale che agisce su di loro.

Dal punto di vista della personalità un aspetto psicologico ampiamente studiato è il desiderio di provare sensazioni forti, nuove ed eccitanti; desiderio che porta i giovani a ricercare esperienze che possano fornire questo vissuto. Questo aspetto, definito *sensation seeking*, è stato misurato in numerosi studi sui comportamenti a rischio mediante la *Sensation Seeking Scale* di Zuckerman (1979). Il punteggio di ricerca di sensazioni è risultato relato all'uso di alcool (Arnett & Balle-Jensen, 1993; Greene, Krcmar, Walters, Rubin & Hale, 2000; Hampson, Severson, Burns, Slovic & Fisher, 2001; Newcomb & McGee, 1989; Schwartz, Burkhart & Grenn, 1978), alla guida in stato di ebbrezza (Arnett, 1990; Arnett & Balle-Jensen, 1993; Greene, Krcmar, Walters, Rubin, Hale, 2000; Zuckerman & Neeb, 1980), all'uso di sostanze stupefacenti (Ball, 1995; Greene, Krcmar, Walters, Rubin, Hale, 2000) ed al comportamento sessuale non protetto (Arnett, 1990b; Arnett & Balle-Jensen, 1993; Greene, Krcmar, Walters, Rubin, Hale, 2000; Sheer & Cline, 1994).

Il problema della componente morale nella percezione del rischio è stato sollevato dagli studi di Sjoberg e Winroth (1986) e Sjoberg e Torell (1993), partendo da una critica allo studio di Benthin, Slovic e Severson (1993), dove mancava completamente l'attenzione a questi aspetti. Negli studi citati gli autori hanno riscontrato una buona correlazione tra la valutazione morale di taluni comportamenti a rischio e l'accettabilità degli stessi. In seguito Savadori e Rumiati (1996) hanno cercato una relazione tra la sfera morale e la percezione del rischio (parzialmente distinguibile dalla accettabilità), riscontrando effettivamente un rapporto molto stretto tra le due: una attività considerata immorale è anche valutata come molto rischiosa e viceversa. In particolare sarebbero proprio i genitori a detenere ed esercitare le norme morali più sentite dai giovani: ragazzi che agiscono comportamenti a rischio ritengono che i

genitori non reagirebbero troppo male, viceversa coloro che non ne adottano ritengono la potenziale reazione dei genitori molto severa. Sul versante sociale, invece, la regolamentazione da parte delle istituzioni non viene percepita dai giovani come efficace (Savadori & Rumiati, 1996).

Infine, dal punto di vista sociale, un fattore fondamentale è l'influenza dei pari, studiata e riconosciuta da molti studi (Benthin, Slovic & Severson, 1993; Clasen & Brown, 1985; Jessor & Jessor, 1975; Savadori & Rumiati, 1996). I giovani si sentono spinti dai coetanei ad adottare comportamenti a rischio ed infatti riferiscono una difficoltà ad evitare di ingaggiare questo tipo di comportamenti (Benthin, Slovic & Severson, 1993). Allo stesso tempo l'influenza dei pari li porta a sentirsi fortemente supportati dai coetanei ed a sopravvalutare la frequenza con cui i coetanei adottano quel tipo di comportamento a rischio. Nel complesso l'influenza dei pari è risultata un fattore determinante nell'uso di alcool (Dishion & Loeber, 1985), nel tabagismo (Chassin, 1985), nell'uso di droga (Galambos & Silbereisen, 1987) e nei comportamenti sessuali (Collins & Harper, 1985).

Esistono infine alcune teorie, basate sui fattori cognitivi o motivazionali della presa di decisione, finalizzate a spiegare il legame tra atteggiamento e comportamento nella assunzione di rischio in adolescenza. Di particolare interesse sono: la Teoria dell'Azione Ragionata (*Theory of Reasoned Action*) di Fishbein e Ajzen (1975) e la Teoria della Auto-Regolazione (*Theory of Self-Regulation*) di Bagozzi (1992).

Lo scopo principale della Teoria dell'Azione Ragionata (Figura 1.1) è quello di comprendere e predire le cause del comportamento basandosi sull'assunto che l'uomo è razionale e quindi prendendo in considerazione fattori prettamente cognitivi (Ajzen & Fishbein, 1980). Il predittore diretto e migliore del comportamento è l'intenzione, vista come punto di transizione tra le componenti cognitive e valutative dell'atteggiamento ed il comportamento vero e proprio (Ajzen, 1988). L'intenzione è, a sua volta, determinata

dall'atteggiamento verso l'azione (una valutazione personale) e dalle norme soggettive (una valutazione sociale). Sia l'atteggiamento che le norme soggettive sono basate su strategie ed informazioni cognitive (Ajzen, 1988).

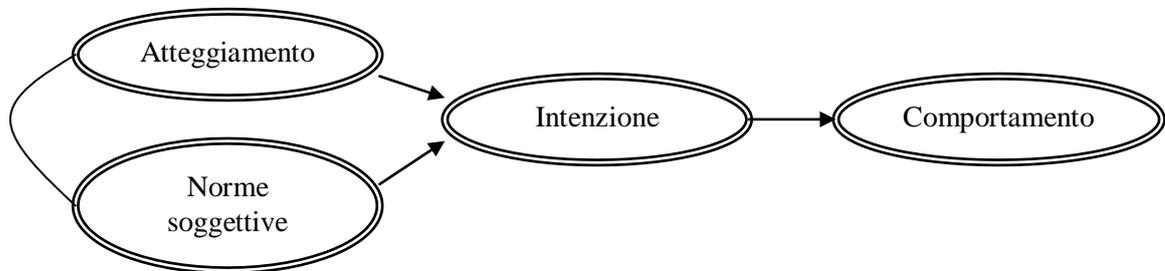


Figura 1.1. Teoria dell'Azione Ragionata. Tratta da Leone, Perugini e Ercolani (1999)

Questo modello è stato proposto dai suoi autori (Ajzen & Fishbein, 1980) come una rappresentazione completa e sufficiente della relazione tra atteggiamento e comportamento, ma gli stessi (Fishbein & Ajzen, 1975) hanno anche precisato che è applicabile solo ai comportamenti volitivi, cioè quei comportamenti per i quali nessun impedimento esterno o interno al soggetto può evitare la messa in atto, una volta che l'intenzione lo abbia stabilito.

La Teoria dell'Azione Ragionata è stata ampiamente utilizzata per spiegare comportamenti quali il tabagismo, l'uso di alcool, l'uso di sostanze stupefacenti ed i comportamenti sessuali non protetti (Fishbein, Middlestadt & Hitchcock, 1994).

Una critica mossa da più parti a questo modello è la sua presunta completezza. Molti studi (Bagozzi, 1981; Bentler & Speckart, 1979, 1981; Fredricks & Dosset, 1983) hanno infatti dimostrato che un fattore, che manca in questo approccio ma è importante per determinare sia l'intenzione che il comportamento presente, è l'esperienza passata, ovvero la frequenza e gli esiti di quello stesso comportamento agito in passato.

La Teoria della Auto-Regolazione (Bagozzi, 1992) è un ampliamento della Teoria dell'Azione Ragionata e non si limita a considerare i fattori cognitivi, bensì introduce anche fattori motivazionali. Il cambiamento di maggiore

rilevanza è l'inclusione del fattore desiderio, che influisce sull'intenzione ed è basato, a sua volta, sulla motivazione (Figura 1.2). Il desiderio è considerato una causa prossimale della intenzione, mentre l'atteggiamento è visto come una causa distale la cui influenza è totalmente mediata dal desiderio. Rimane l'influenza delle norme soggettive sulle intenzioni e delle intenzioni sul comportamento.

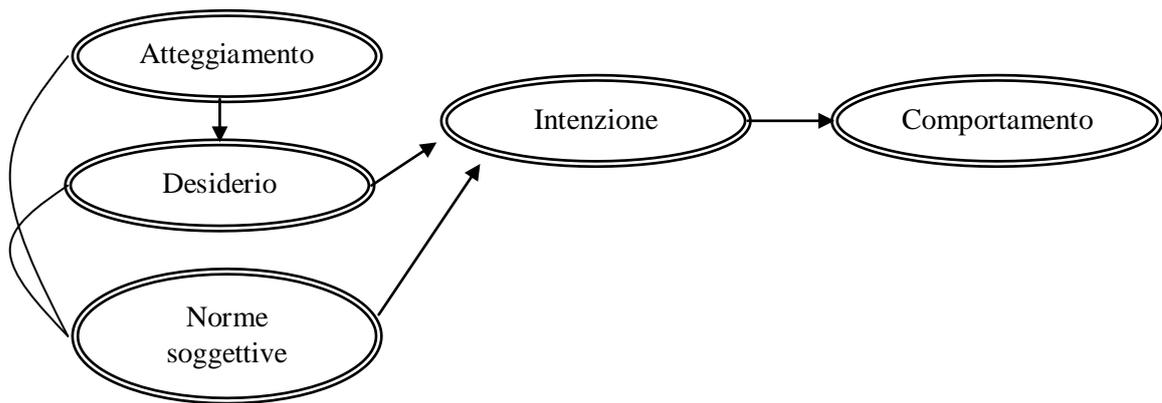


Figura 1.2. Teoria della Auto-Regolazione. Tratta da Leone, Perugini e Ercolani (1999)

In questo modello le intenzioni continuano ad essere un buon predittore del comportamento, ma Bagozzi (1992) include anche l'influenza del comportamento passato che viene suddivisa in due componenti: la frequenza del comportamento e la *recency*, cioè quanto esso è avvenuto di recente (Bagozzi & Warshaw, 1990). Questi due aspetti sono considerati concettualmente distinti e gli autori ipotizzano che influenzino l'intenzione ed il comportamento presente in modi e con intensità diversi.

La Teoria dell'Azione Ragionata e la Teoria della Auto-Regolazione sono state messe a confronto, per quanto concerne il loro potere predittivo del comportamento in adolescenza, in un interessante studio di Leone, Perugini e Ercolani (1999). In realtà lo studio prende in considerazione anche la Teoria del Comportamento Pianificato (Ajzen, 1985), considerandola come un passaggio intermedio tra le altre due. Dal confronto emerge chiaramente la superiorità predittiva della Teoria della Auto-Regolazione rispetto alla Teoria dell'Azione

Ragionata e l'influenza del comportamento passato, soprattutto nei termini di *recency* (Leone, Perugini & Ercolani, 1999).

Infine i fattori motivazionali sono alla base di una delle teorie più utilizzate nella spiegazione del comportamento a rischio in adolescenza, la già citata Teoria del Comportamento Problematico (*Problem Behavior Theory*) di Jessor e Jessor (1977). Secondo questo approccio gli adolescenti sarebbero motivati ad adottare comportamenti a rischio per ottenere degli obiettivi specifici, che ritengono di non poter raggiungere altrimenti: dimostrarsi adulti, affrancarsi dal controllo parentale, affermare la propria indipendenza, essere ammessi nel gruppo dei pari.

Capitolo 2

La Teoria della Detezione del Segnale

2.1 Introduzione

La Teoria della Detezione del Segnale (TDS) è stata originariamente sviluppata nell'ambito della psicofisica, ma ha trovato in seguito ampia applicazione nella ricerca in psicologia poiché la si può utilizzare proficuamente nell'analisi di due tipi di processi di interesse psicologico: la detezione ed il riconoscimento. Un qualsiasi compito di diagnosi, inteso nella sua accezione più ampia, può coinvolgere un processo di detezione, riconoscimento o entrambi; e risulta chiaro quanto l'analisi di questo tipo di procedimento sia particolarmente interessante in vari ambiti della psicologia e non solo.

Un compito di detezione ha lo scopo di determinare se uno stimolo, di una specifica categoria, è presente o meno; mentre un compito di riconoscimento stabilisce se lo stimolo presentato fa parte della categoria A o della categoria B. Inoltre questi due tipi di compiti coinvolgono distinti processi cognitivi: la discriminazione e la decisione. La prima stabilisce il grado in cui l'evidenza in una osservazione va a favore dell'occorrenza dell'evento; mentre il processo di decisione determina quanto l'evidenza debba essere forte in favore all'alternativa A affinché questa sia proprio l'alternativa scelta. In entrambi i casi, la forza dell'evidenza si situa lungo un continuum probabilistico ed il decisore stabilisce un punto di scelta – o *cut-off* – lungo questo continuum: questa soglia viene definita nella TDS come criterio di decisione; ed è proprio questo aspetto della teoria che è stato studiato, approfondito ed applicato nel presente lavoro.

In altre parole la discriminazione (definita anche accuratezza o sensibilità) riflette l'abilità dell'osservatore di distinguere tra le classi di stimoli, mentre il criterio di decisione (detto anche *response bias*) rappresenta la tendenza a scegliere una risposta piuttosto che un'altra.

La TDS raccomanda che i comportamenti, osservati durante i compiti, vengano scomposti in modo che la discriminazione e il criterio di decisione possano essere valutati in modo indipendente.

Prima di entrare nel merito della TDS e specificamente dell'analisi ROC (originariamente *Receiver Operating Characteristic*, ridefinito da Swets nel 1996 come *Relative Operating Characteristic*), vogliamo ripercorrere brevemente la sua storia dalle origini agli sviluppi più recenti.

2.2 La Storia della TDS

Come già accennato, la TDS nasce nell'ambito della **psicofisica** e dobbiamo quindi partire dal capostipite di questa disciplina: Gustav Fechner. Questo studioso intorno al 1860 ha ideato quelli che sono poi divenuti i metodi della psicofisica di base e li ha utilizzati per misurare sia la minore differenza registrabile tra due stimoli (soglia differenziale) sia la minore intensità di uno stimolo possibile da esperire (soglia assoluta). In termini moderni possiamo ridefinire il compito di soglia assoluta come detezione e quello di soglia differenziale come riconoscimento. La detezione può essere considerata anche un caso speciale di riconoscimento quando uno dei due stimoli è nullo.

Lo stesso Fechner ha messo in evidenza che il primo problema con cui i metodi psicofisici si devono confrontare è la grande variabilità dovuta alle differenze individuali, al tempo e ad altre innumerevoli condizioni interne ed esterne che influenzano il processo di decisione. Come è noto: replicare molte volte le prove è l'unico modo per contravvenire a questi errori. Inoltre Fechner è stato il primo a mettere graficamente in relazione la proporzione di risposte positive con la misura della forza dello stimolo (o della differenza tra stimoli) per creare una funzione psicometrica.

Il secondo importante esponente della psicofisica che ha fatto enormi passi avanti nello studio della discriminazione è stato Louis Leon Thurstone negli anni 20 del XX secolo. Il suo modello parte dalla rappresentazione di due

distribuzioni sovrapposte che rappresentano la grandezza di due stimoli simili (Figura 2.1). Il modello include inoltre degli assunti molto specifici: le distribuzioni devono essere normali, con uguale deviazione standard e con una correlazione tra gli stimoli pari a zero. E' inoltre possibile calcolare la differenza tra le medie delle due distribuzioni: questa, denominata d , è espressa in deviazioni standard.

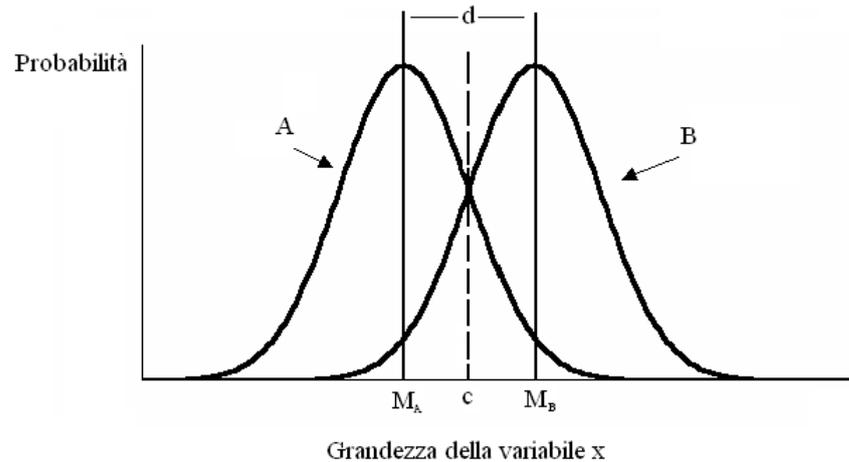


Figura 2.1. Distribuzioni delle funzioni di densità di probabilità. Modello di Thurstone.

Thurstone utilizza come solo parametro proprio la differenza d , poiché parte dall'assunto che la procedura a scelta forzata elimini ogni tipo di errore da parte dell'osservatore: in altre parole l'osservatore parte dall'idea che entrambi gli eventi siano ugualmente probabili e desiderabili.

Se si definisce l'errore di risposta in termini di criterio di decisione, questo criterio è rappresentato da un punto di *cut-off* (c) lungo l'asse di grandezza della variabile osservata (x), tale per cui valori di x inferiori a c determinano la risposta A mentre valori di x maggiori di c determinano la risposta B. L'assunto di Thurstone, relativamente all'uguaglianza nella probabilità a priori e nella desiderabilità degli eventi, equivale a porre il criterio di decisione nel punto in cui le due distribuzioni si incrociano.

Come vedremo meglio in seguito, l'analisi ROC permette invece all'osservatore di posizionare il criterio di decisione ovunque nell'intero range di

valori di x e di estrarre una misura sia del criterio che di discriminazione, indipendenti tra loro.

Negli anni 40 del XX secolo, H. Richard Blackwell (1952) ha focalizzato i propri studi sui compiti di detezione in cui uno degli stimoli è nullo. In realtà, nel periodo in cui Blackwell lavorava in questo ambito, la concezione di stimolo nullo è cambiata: esso è stato inizialmente considerato appunto nullo, mentre in seguito lo si considera uno stimolo esso stesso. Questa nuova concezione include nel processo di decisione anche una quota di variabilità, attribuita sia al procedimento che al partecipante, definita “rumore” che può produrre una attivazione sensoriale e neurale confondibile con l’effetto elicitato dallo stimolo vero e proprio. In questa nuova concezione le distribuzioni di Thurstone rappresentano rispettivamente: la distribuzione di probabilità del “rumore” (n da *noise*), e la distribuzione data dal segnale più il rumore ($s+n$ da *signal+noise*).

Anche se Blackwell (1952) aveva compreso che il rumore poteva interferire con la detezione del segnale, fece una ulteriore assunzione che, di fatto, eliminò ogni possibilità che il rumore fosse preso per segnale. Assunse l’esistenza di un criterio di decisione per le risposte positive ad un livello tale per cui raramente un osservatore avrebbe potuto essere confuso dal rumore riportando un segnale (Figura 2.2).

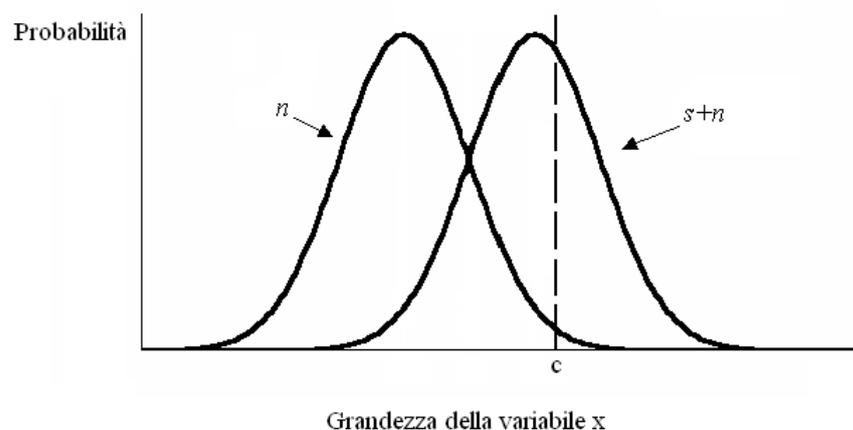


Figura 2.2. Distribuzioni delle funzioni di densità di probabilità. Modello di Blackwell.

Come mostrato nella Figura 2.2, questo criterio ricorda molto da vicino la regola classica per la verifica della significatività dell'ipotesi nulla in statistica ad un livello di confidenza dell'1% ($\alpha=0.01$). Blackwell (1952) assunse inoltre che valori dello stimolo inferiori al criterio c non fossero distinguibili e che quindi sostanzialmente l'osservatore potesse solo indovinare. In questo senso infatti ideò un metodo di correzione per *guessing* che vedremo nel dettaglio nel paragrafo relativo agli indici di accuratezza. È importante sottolineare che questa correzione parte dall'assunto che risposte positive giuste ed errate siano statisticamente indipendenti, mentre i risultati successivi hanno ampiamente dimostrato che questo assunto non è corretto.

Uno dei contributi fondamentali che hanno reso la TDS una teoria moderna è stato quello della **teoria statistica della decisione**. Questa ultima ha anch'essa come punto di partenza la rappresentazione grafica di due distribuzioni sovrapposte (Figura 2.3).

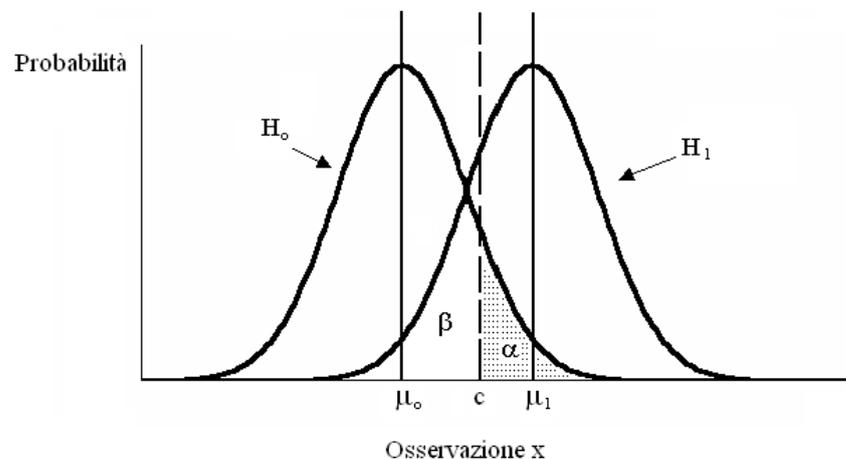


Figura 2.3. Distribuzioni delle funzioni di densità di probabilità. Teoria statistica della decisione.

Sull'asse orizzontale si trova la variabile x , che è una misura della forza dell'osservazione o della grandezza della statistica campionaria. Sull'asse verticale si trova invece la densità di probabilità. La distribuzione di sinistra è la funzione di densità di probabilità di x per l'ipotesi nulla H_0 , quella di destra è la funzione di densità di probabilità per l'ipotesi alternativa H_1 . Il grado di

sovrapposizione delle distribuzioni determina il grado di incertezza nel discriminare tra le due alternative. Thurstone (1927) nella teoria psicometrica denominò queste distribuzioni “dispersioni discriminanti”.

La costruzione di un test statistico equivale a dividere l’asse delle x in due regioni ponendo un criterio di decisione (c) di modo che valori campionari di x inferiori a c portino ad accettare H_0 , e valori maggiori di c inducano ad accettare H_1 . La posizione del criterio determina la probabilità associata a due possibili tipi di errori: errore di primo tipo o α (accettare H_1 quando è vera H_0) ed errore di secondo tipo o β (accettare H_0 quando è vera H_1). In generale si cerca di aggiustare queste due probabilità basandosi sui costi connessi ai due possibili tipi di errore. Solitamente la strategia è quella di fissare arbitrariamente la probabilità dell’errore di primo tipo (di solito a .05 o .01) e poi scegliere il criterio che minimizza la probabilità dell’errore di secondo tipo.

Jerzy Neyman e Egon Pearson (1933) hanno definito le regole più importanti di questa procedura di verifica e hanno introdotto un termine fondamentale: il rapporto di verosimiglianza (*likelihood ratio*) dato dal rapporto tra il valore di ordinata di H_1 e il valore di ordinata di H_0 per un certo valore di x . Si accetta H_1 quando il rapporto supera il valore c , laddove c è stato scelto a priori per determinare la probabilità desiderata associata all’errore di primo tipo. Solitamente, piuttosto che considerare la probabilità dell’errore di secondo tipo, si dà maggiore importanza alla potenza del test data da $(1-\beta)$, ovvero la probabilità di accettare H_1 quando è vera. Quindi, secondo Neyman e Pearson (1933), è necessario fissare la probabilità dell’errore di primo tipo e poi scegliere un rapporto uguale a c in modo da massimizzare la potenza del test.

Una differenza tra questo approccio statistico e l’analisi ROC sta nel fatto che quest’ultima conferisce uguale importanza ai due tipi di errore e mostra come essi covariano al variare del criterio per ogni possibile differenza tra le medie delle due distribuzioni.

Negli anni 40 Abraham Wald (1950) ha costruito una serie di modelli relativi a strategie di decisione ottimali in situazioni in cui la fonte di evidenza fosse complessa ed incerta. Quando la distribuzione dell'evidenza è nota, la decisione statistica ottimale è una soluzione Bayesiana che confronti le probabilità che ciascuna decisione alternativa sia vera o valida: la variabile di decisione ottimale è il rapporto tra le probabilità e la scelta migliore è quella che favorisce l'alternativa avente maggiore probabilità. La sfida di Wald (1950) è stata quella di sviluppare una teoria della decisione statistica per quelle situazioni in cui non è possibile conoscere a priori la distribuzione e quindi applicare il modello Bayesiano. Wald (1950) ha dimostrato che le strategie di decisione ottimali si basano comunque sulla stima della verosimiglianza delle alternative basandosi sull'evidenza a disposizione. In tutti i casi in cui non sia possibile ipotizzare o conoscere a priori la distribuzione dell'evidenza, la migliore variabile statistica di decisione è il rapporto di verosimiglianza.

2.3 La Teoria della Detezione del Segnale e l'Analisi ROC

Grazie ai contributi delle varie teorie, intorno alla metà del secolo scorso è nata la TDS e con essa una serie di procedure di analisi che tra breve verranno illustrate.

Prima di entrare nel merito dell'analisi ROC è necessario chiarire un aspetto fondamentale della TDS: la variabile oggetto di decisione. Un problema fondamentale nella teoria della decisione statistica è la natura della variabile di decisione e la TDS mutua questa problematica e la sua risoluzione proprio dalla teoria statistica. In quest'ultima la variabile su cui viene presa la decisione è solitamente una funzione matematica derivata dai dati osservati, ed è proprio la probabilità connessa a questa variabile che viene valutata per decidere. Nella TDS avviene un procedimento sostanzialmente identico. L'osservatore compie il processo di scelta in uno spazio multidimensionale, laddove ciascuna osservazione deriva da molte e diverse dimensioni. Per ciascuna di queste

dimensioni l'osservatore compie una valutazione in merito alla probabilità che quella caratteristica emerga da una determinata alternativa di scelta tra le possibili. Tutte queste valutazioni vengono combinate insieme a formare un giudizio unitario riguardo alla alternativa più probabile: il rapporto di verosimiglianza (Pastore, Crawley, Berens & Skelly, 2003). Questa nuova variabile, data dalla combinazione delle probabilità dei vari eventi presenti nello spazio multidimensionale, è unidimensionale e diviene la variabile oggetto della decisione. La natura unidimensionale della variabile di decisione da un lato favorisce senza dubbio le possibilità di calcolo, tuttavia questo non significa che la TDS sia incapace di descrivere statisticamente fonti di evidenza anche molto complesse e multidimensionali (Pastore, Crawley, Berens & Skelly, 2003).

Tornando all'oggetto del presente paragrafo, la TDS ha messo a punto una procedura di analisi particolarmente proficua nella verifica dei sistemi diagnostici di qualsiasi tipo: la analisi ROC. Essa è stata applicata sia ad una serie di ambiti strettamente legati alle funzioni percettive (processi sensoriali, vigilanza, selettività percettiva e memoria) sia ad ambiti cognitivi (attenzione, apprendimento, giudizio concettuale, linguaggio) sia, infine, alla valutazione di processi più complessi, quali la presa di decisione in ambito medico o la funzionalità dei sistemi di recupero delle informazioni.

Il processo di discriminazione è basato su considerazioni, quali la probabilità e l'utilità, che sono estranee e potenzialmente confondenti rispetto alla semplice misura della discriminazione di per sé. L'analisi ROC è una tecnica, derivante dalla teoria della decisione statistica e dalla teoria della detezione elettronica, che riesce ad isolare efficacemente gli effetti del criterio di decisione dell'operatore da quelli del comportamento di discriminazione. In molti ambiti di studio della psicologia, effetti che si pensava fossero legati alle proprietà del processo di discriminazione si sono rivelati, attraverso l'analisi ROC, legati piuttosto alle caratteristiche del processo di decisione ed al criterio ad esso connesso.

Negli anni 50 Peterson, Birsdall e Fox (1954) svilupparono l'analisi ROC indipendentemente e parallelamente rispetto a Tanner e Swets (1954a), ma mostrarono per primi la rappresentazione grafica dei dati nota come curva ROC o curva *isosensitiva*, cioè tracciata a sensibilità costante (Figura 2.4).

Quest'ultima è un grafico costruito facendo variare il criterio di decisione e ponendo in ordinata la potenza del test ($1-\beta$ o probabilità di rispondere correttamente) e in ascissa la probabilità dell'errore di primo tipo (α o probabilità di commettere un falso allarme).

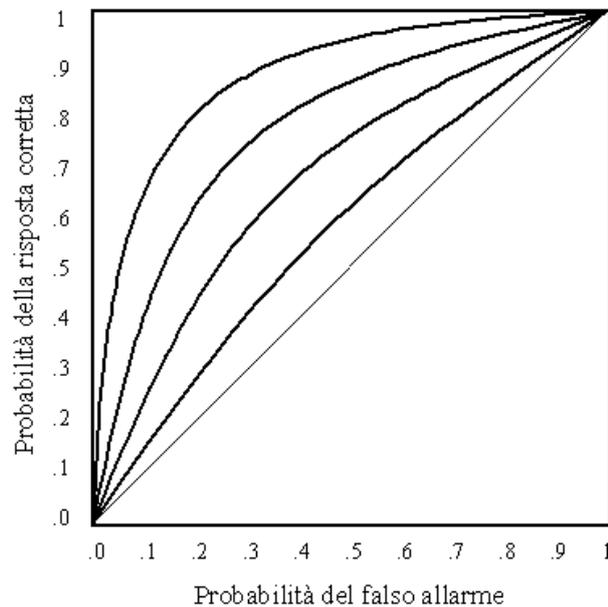


Figura 2.4. Curve ROC tracciate variando l'accuratezza del sistema.

L'analisi ROC fornisce una misura di discriminazione (o accuratezza) che è indipendente dalla posizione del criterio di decisione e non influenzata dai processi, ad esempio di aspettativa e motivazione, che influiscono sulla risposta. Questo parametro è dato dalla differenza tra le medie delle due distribuzioni ed è denominato d quando è espresso in varianza, mentre, se misurato in deviazioni standard, è denominato d' , dove $d' = \sqrt{d}$.

Allo stesso tempo però l'analisi ROC fornisce una misura dell'effetto dei processi, inerenti l'osservatore, che influiscono sulla risposta: essa è data dal

cosiddetto criterio di decisione (o *response bias*). La misura del criterio di decisione tipica del modello TDS è chiamata β , ed è data dal valore del rapporto di verosimiglianza (*likelihood ratio*) a cui il criterio deve essere posizionato per dare un punto particolare sulla curva ROC. In altre parole β è uguale al rapporto tra le ordinate corrispondenti ai punti in cui una perpendicolare all'ascissa innalzata dal punto c interseca le due distribuzioni (Figura 2.5):

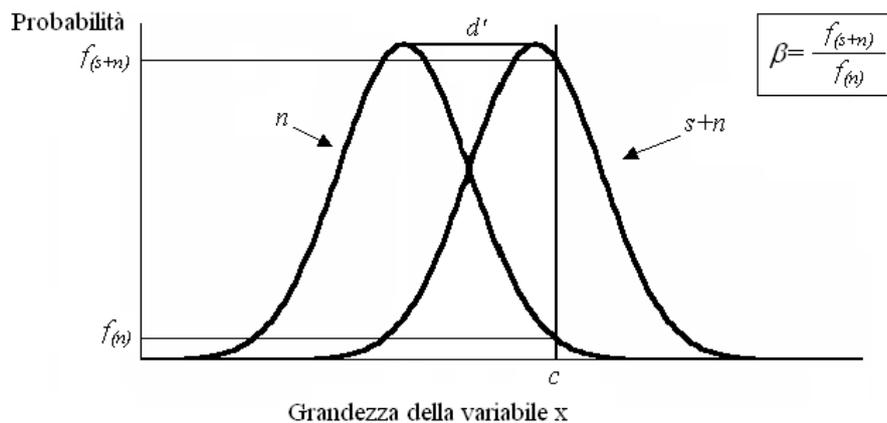


Figura 2.5. Rappresentazione grafica del calcolo del criterio di decisione β in funzione delle distribuzioni di densità di probabilità.

L'indice di accuratezza d' e quello relativo al criterio di decisione β sono le misure classiche della TDS, ma presentano dei vincoli importanti: il modello prevede che le distribuzioni di probabilità di $(s+n)$ e (n) siano normali ed aventi uguale varianza. La normalità delle distribuzioni è, peraltro, un requisito fondamentale per garantire l'indipendenza dei due parametri.

E' necessario aprire una parentesi e fare un po' di chiarezza su cosa debba intendersi per indipendenza tra d' e β , poiché spesso vi sono fraintendimenti in merito. Indipendenza significa primariamente che il valore del criterio di decisione non deve influire sull'accuratezza, in altre parole il valore di d' deve riflettere solo la capacità di discriminazione determinata dall'evidenza e deve essere costante al variare di β (Pastore, Crawley, Berens & Skelly, 2003). Questo è l'aspetto fondamentale dell'indipendenza tra questi due parametri.

Il modello gaussiano di TDS è senza dubbio quello più utilizzato, ma non è da considerarsi l'unico possibile ed è comunque necessario fare molta attenzione ad utilizzare dei parametri quali d' e β senza aver preventivamente controllato che sussistano le condizioni per il loro utilizzo (Pastore, Crawley, Berens & Skelly, 2003). Per tutte quelle situazioni in cui questi assunti vengano violati, sono stati creati ed utilizzati altri indici, che vedremo nello specifico in seguito, alcuni dei quali presuppongono comunque dei vincoli, mentre altri sono supposti essere non parametrici.

Tornando alla curva ROC, essa può essere visualizzata in un grafico "binormale": in cui le probabilità in coordinate vengono espresse mediante delle scale di misura in cui i corrispondenti valori di deviazione sono disposti in modo da aumentare linearmente con un rapporto costante (*linearly scaled*). In altre parole, questo significa che le coordinate del grafico sono trasformate ed espresse in punti z e vengono disposte in modo corrispondente sui due assi. In questo grafico la ROC è rappresentata da una linea retta ad inclinazione variabile: solitamente tra 0.5 e 1.5 (Figura 2.6).

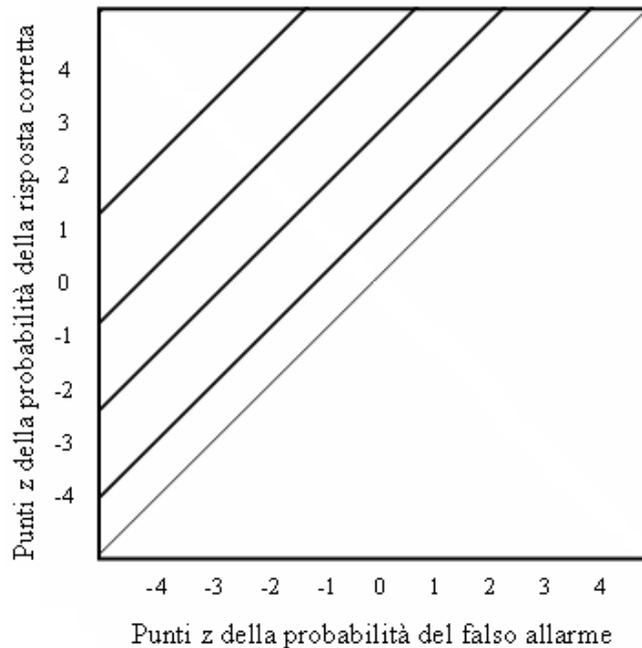


Figura 2.6. Curve ROC su grafico binormale tracciate variando l'accuratezza del sistema.

Un modo diretto per tracciare una ROC è istruire l'osservatore ad adottare diversi criteri di decisione da una prova all'altra, in modo da ottenere numerosi punti nello spazio ROC. Un modo però più efficiente, e raccomandato da Swets (1996), è far utilizzare all'osservatore una *rating scale* in un unico gruppo di prove: ad esempio può essere chiesto all'osservatore di segnare su una scala a cinque punti la fiducia che ripone nella correttezza della propria risposta. E' inoltre buona norma costruire una ROC basandosi sulle risposte di numerosi osservatori.

Swets (1996) passa in rassegna le caratteristiche di una serie di ROC empiriche, ottenute sia in vari ambiti della psicologia sperimentale (visione, memoria di riconoscimento sia per parole che per odori, apprendimento animale e giudizio concettuale) sia in ambiti pratici (radiologia, recupero delle informazioni, meteorologia, verifica attitudinale, verifica della veridicità di affermazioni tramite poligrafo). La sua conclusione è che tutte le curve ROC utilizzate sono risultate opportunamente costruite mediante il grafico binormale sotto forma di linee rette ad inclinazione variabile.

I dati raccolti in un compito di discriminazione, del tipo di quelli discussi in precedenza, possono essere rappresentati, per una certa posizione del criterio di decisione, da una tavola di contingenza 2x2 avente come variabili stimoli e risposte e come contenuto le frequenze dei quattro possibili esiti (Tabella 2.1).

Tabella 2.1
Matrice di decisione statistica

Scelta	Evento	
	A	B
A	Risposta corretta o Hit	Falso allarme o Errore di primo tipo - α
B	Omissione o Errore di secondo tipo - β	Rifiuto corretto

All'interno di questa tabella i dati rilevanti ai fini dell'analisi ROC sono: la probabilità condizionale del falso allarme e la probabilità condizionale della

risposta corretta o hit. Infatti, come abbiamo già visto, la ROC è un grafico che mostra la probabilità condizionale di scegliere l'alternativa A quando questa è presente (hit o h) in funzione della probabilità condizionale di scegliere A quando è presente B (falso allarme o f). Sia h che f crescono all'aumentare della tendenza a scegliere A, ovvero via via che il criterio di scelta di A diviene più indulgente.

Nei paragrafi seguenti saranno presentati gli indici di accuratezza della discriminazione e le misure del criterio di decisione maggiormente utilizzati nella analisi ROC. E' necessario però fare una breve premessa a questa trattazione.

Gli indici classici della TDS, ad esempio d' e β , hanno dei vincoli ben precisi che riguardano la forma delle distribuzioni implicate e dei vantaggi che consistono nella possibilità di essere calcolabili a partire da una singola coppia di valori di probabilità di hit e falsi allarmi. Il problema dell'assunto sulla forma delle distribuzioni è certamente un limite nelle possibilità di applicazione dei modelli di detezione, dal momento che talvolta non è possibile determinare la distribuzione di probabilità degli eventi. Alcuni studiosi (Green, 1964; Green & Swets, 1966; Grier, 1971; Hodos, 1970; Pollack & Norman, 1964) hanno quindi cercato di mettere a punto delle misure che fossero indipendenti dalle distribuzioni e che hanno definito come non parametriche, *distribution free* o anche *triangle-based* (cioè basate sui triangoli dello spazio ROC).

Alcune critiche sollevate negli anni seguenti (McNicol, 1972) non sono state adeguatamente valutate e queste misure sono state accettate e utilizzate ad ampio raggio, fino a che, dopo circa un ventennio di fama indiscussa, alcuni autorevoli studiosi (Macmillan & Creelman 1990, 1996; Pastore, Crawley, Berens & Skelly, 2003; Snodgrass & Corwin, 1988; Swets, 1986) hanno messo fortemente in discussione e scardinato la supposta non parametricità della maggior parte di queste misure, dimostrandone l'equivalenza o la dipendenza

matematica da altri parametri che sottendono specifiche distribuzioni di probabilità.

Nella discussione relativa alle misure, sia di accuratezza che di decisione, questo tipo di indici verranno presi in considerazione dandone per assodata la non parametricità, evitando di entrare oltre nel merito della discussione relativa al problema della parametricità, poiché offrirne una trattazione completa non rientra tra gli obiettivi del presente lavoro.

2.3.1 Misure di accuratezza

Swets (1996) sottolinea più volte che il modo migliore di quantificare l'accuratezza è farlo indipendentemente dalla frequenza relativa degli eventi e dal criterio di decisione.

I sistemi diagnostici, di qualsiasi tipo siano, sono solitamente chiamati a discriminare tra due alternative possibili. Come abbiamo già detto, avendo a disposizione due possibili eventi e due corrispondenti tipi di diagnosi, i dati che si ottengono sono quelli di una tavola di contingenza 2x2 (Tabella 2.2).

Tabella 2.2
Tavola di contingenza 2x2:
frequenze di cella, di riga, di colonna e totali

Scelta	Evento		
	A	B	
A	a	b	a+b
B	c	d	c+d
	a+c	b+d	N

Il rapporto tra la probabilità di una cella e la somma della sua colonna rappresenta una probabilità condizionale: la probabilità di scelta condizionata dalla reale occorrenza dell'evento. Come già sappiamo, le probabilità condizionate che maggiormente interessano l'analisi ROC sono quella relativa alla risposta corretta (hit o h) e quella relativa al falso allarme (f); che possono essere calcolate mediante le formule 2.1 e 2.2:

$$h = \frac{a}{a + c} \quad (2.1)$$

$$f = \frac{b}{b + d} \quad (2.2)$$

Le ROC possono assumere varie forme, ma è possibile raggrupparle in tre distinte categorie che prevedono modelli e indici di accuratezza peculiari: 1) indici basati su distribuzioni rettangolari o modelli a soglia fissa; 2) indici basati su distribuzioni continue con varianze omogenee o modelli a criterio variabile; 3) indici basati su distribuzioni continue e rettangolari con varianze non omogenee o modelli a criterio variabile (Tabella 2.3).

Gli indici delle prime due categorie possono essere ottenuti partendo dalle frequenze della tabella 2x2, richiedono specifici assunti in merito alla forma delle distribuzioni sottostanti ma, allo stesso tempo, possono essere calcolati a partire da una singola coppia di valori di probabilità di hit e falsi allarmi. Per gli indici della terza categoria il calcolo si basa direttamente sulla curva ROC e, laddove il vantaggio è quello di non richiedere assunti in merito alle distribuzioni, il limite è che per garantire la non parametricità è necessario avere a disposizione l'intera curva ROC e non è sufficiente una coppia di valori.

Gli indici basati su distribuzioni rettangolari predicono ROC binormali curvilinee. I primi due indici di questa categoria (H_c e H'_c) sono versioni della probabilità di hit corretta per *guessing* (Tabella 2.3). Il primo è utilizzato primariamente negli studi sulle funzioni sensoriali, il secondo, invece, ha trovato utilizzo negli studi sulla memoria di riconoscimento, in meteorologia e diagnosi medica. Entrambi sono centrati su h , ma cercano di depurare questa misura da componenti spurie che possono indurre una tendenza verso il falso allarme o f .

Un ulteriore e molto utilizzato indice di questa categoria (Tabella 2.3) è la proporzione di risposte corrette $P(C)$. Rappresenta la percentuale di risposte corrette di entrambe le alternative ed è utilizzato solitamente in meteorologia e diagnosi medica. E' un indice molto discusso e della cui utilità Swets (1996)

dubita fortemente, fino a sconsigliarne l'utilizzo, pur facendo una distinzione: nel caso di un compito (asimmetrico) di scelta di una alternativa (sì-no), l'indice $P(C)$ risulta povero; mentre in un compito (simmetrico) di scelta tra due

Tabella 2.3
Classificazione delle misure di accuratezza

Indici basati su distribuzioni rettangolari		
Probabilità di hit corretta per <i>guessing</i> (Blackwell, 1963; Fisk & Schneider, 1984)	$H_c = \frac{(h-f)}{(1-f)}$	<i>high threshold model</i>
	$H'_c = h - f$ (Gillund & Shiffrin, 1984; Woodworth, 1938)	
Proporzione di risposte corrette (Finley, 1884)	$P(C) = \frac{(a+d)}{N}$	<i>double threshold model</i>
Skill test (Woodcock, 1976)	$Z = \frac{4(ad-bc)}{N^2}$	
Statistica kappa (Landis & Koch, 1977)	$K = \frac{2(ad-bc)}{2(ad-bc) + N(b+c)}$	
Coefficiente di correlazione punto-tetracorico (Hays, 1973)	$\phi = \frac{(ad-bc)}{[(a+c)(b+d)(a+b)(c+d)]^{\frac{1}{2}}}$	
Indici basati su distribuzioni continue con varianze omogenee		
d' (Tanner & Swets, 1954b)	$d' = z_f - z_h$	distribuzione normale
Indice η della Teoria della Scelta (Luce, 1959; 1963)	$\eta = \left(\frac{bc}{ad}\right)^{\frac{1}{2}}$	
Indice Log Odds Ratio (Goodman, 1970)	$LOR = \ln\left(\frac{ad}{bc}\right)$	distribuzione logistica
Indice Q (Yule, 1912)	$Q = \frac{(ad-bc)}{(ad+bc)}$	
Indici basati su distribuzioni continue e rettangolari con varianze non omogenee		
Indice A' (Pollack & Norman, 1964)	$A' = \frac{1}{2} + \frac{1}{4} \frac{(H-F)(1+H-F)}{H(1-F)}$	per $H \geq F$

<p>Indice A_z (Swets & Pickett, 1982)</p>	$A' = \frac{1}{2} + \frac{1}{4} \frac{(F-H)(1+F-H)}{F(1-H)} \quad \text{per } H \leq F$ $A_z = \Phi \left[\frac{\text{Intercetta}}{\sqrt{1 + (\text{Slope}^2)}} \right]$
--	---

alternative (scelta forzata), $P(C)$ può essere considerato teoricamente uguale alla porzione di area sotto la curva ROC ed il suo utilizzo è appropriato.

Gli ultimi tre indici basati su distribuzioni rettangolari (Tabella 2.3) sono: lo *skill test* di Woodcock (1976), nato in meteorologia, misura il grado in cui la discriminazione eccede la performance casuale; la statistica kappa (Landis & Koch, 1977), solitamente utilizzata per misurare l'accordo tra osservatori in medicina clinica, è sostanzialmente l'indice $P(C)$ corretto per *guessing*; il coefficiente di correlazione punto-tetracorico phi (Hays, 1973), utilizzato in psicologia sperimentale, meteorologia e nei procedimenti di verifica dei materiali.

L'indice H_c prevede un modello alta-soglia (*high-threshold model*), quindi con due soli stati: segnale o rumore, mentre gli indici H'_c , $P(C)$, Z , K e ϕ prevedono un modello a doppia soglia (*double threshold model*) e, a parte ϕ , corrispondono ad un modello a tre stati: le due soglie definiscono tre diverse categorie o rappresentazioni delle possibili risposte. Per la teoria della detezione (Macmillan & Creelman, 1990), i modelli a soglia assumono che lo spazio decisionale sia caratterizzato da un numero ristretto di stati discreti, basati su distribuzioni uniformi a forma rettangolare. Da questo punto di vista i modelli a soglia, nonostante facciano parte delle teorie di detezione, si differenziano dalla TDS e dalla Teoria della Scelta che sono, altresì, tra loro piuttosto simili ed ipotizzano una spazio decisionale continuo (Macmillan & Creelman, 1990).

Tra gli indici basati su distribuzioni continue con varianze omogenee, quello in assoluto più famoso ed utilizzato nell'analisi ROC è il già menzionato indice d' (Tanner & Swets, 1954b). Il calcolo di d' parte dai dati della tabella di contingenza e si basa sull'utilizzo della tavola statistica della curva normale.

Ciascuna delle probabilità associate a hit (h) e falsi allarmi (f) è trattata come la porzione di area sottostante la curva normale e, mediante la tavola statistica adatta, è possibile trovare il punto z corrispondente ad entrambi i valori. A questo punto d' è dato dalla formula in Tabella 2.3.

Gli indici seguenti (η , LOR , Q) sono simili a d' , ma dipendono da distribuzioni logistiche, non normali (Tabella 2.3). L'indice d' predice ROC puramente lineari di inclinazione 1, mentre gli altri tre indici implicano ROC leggermente curve, che sono sostanzialmente indistinguibili da quella lineare. Il modello a criterio variabile su cui si basa d' parte dall'assunto che le due distribuzioni siano normali e di uguale varianza, mentre questo stesso modello applicato agli altri tre indici prevede delle distribuzioni logistiche di uguale varianza, simili comunque alla normale.

Infine (Tabella 2.3) gli indici basati su distribuzioni continue e rettangolari con varianze non omogenee derivano dall'area dello spazio sotto la curva ROC ed hanno l'ambizione di poter essere non parametrici, cioè indipendenti dalla forma delle distribuzioni. Le misure basate sull'area nello spazio ROC assumono valori che vanno da 0.5 per una performance casuale ad 1 per una performance perfetta.

La porzione di area è uguale alla proporzione di risposte corrette in un compito a scelta forzata tra due alternative (Green, 1964; Green & Swets, 1966) e proprio questa caratteristica ha contribuito a far considerare questo tipo di indici come misure non parametriche. In vero va precisato che in un compito *rating*, in cui si abbia a disposizione l'intera curva ROC, l'area sotto la stessa è davvero una misura non parametrica, il problema sorge quando si voglia stimare questa misura a partire da un singolo punto, cioè da una singola coppia di valori di probabilità di hit e falsi allarmi (Green & Swets, 1966).

Swets (1996) stesso ha più volte riconosciuto che gli indici derivati dalla tavola di contingenza 2x2 confondono l'abilità nella discriminazione e il criterio di decisione, mentre l'indice di accuratezza ideale dovrebbe essere indipendente

sia dal criterio che dalla forma delle distribuzioni di probabilità. Seguendo lo stesso ragionamento, lo studioso considera molto importanti le misure basate sull'area della ROC, ma ne discute la parametricità, arrivando a ritenere che l'unica misura di accuratezza realmente non parametrica è la porzione di area che si trova sotto la ROC in un grafico binormale, denominata A_z .

Storicamente la prima misura dell'area sotto la curva ROC (Tabella 2.3) è stata l'indice A' di Pollack e Norman (1964), che fu proposto come misura libera di assunti in merito alle distribuzioni e calcolabile a partire da una singola coppia di valori di hit e falsi allarmi. Per un unico punto dello spazio ROC esistono molte possibili curve: l'indice A' è una sorta di media tra le performance migliore e peggiore a partire da questo punto. Macmillan e Creelman (1996) dimostrano che in realtà A' dipende da due misure di accuratezza che appartengono a diversi modelli di detezone: α , il parametro logistico di accuratezza basato sulla differenza tra medie; e $P(C)$, la proporzione di risposte corrette, cioè il parametro di accuratezza per il modello a doppia soglia, quindi avente distribuzioni rettangolari. In altri termini, quando si ha una bassa accuratezza, A' è dominato da α , quindi sottintende una distribuzione logistica, mentre, quando l'accuratezza è alta, è più importante il termine $P(C)$ e le distribuzioni sottostanti saranno rettangolari. In nessuno dei due casi è possibile definire l'indice A' come non parametrico.

Tra gli indici basati su distribuzioni continue e rettangolari con varianze non omogenee, A_z rappresenta, come detto, la porzione di area sotto la ROC in un grafico binormale (Figura 2.7).

Per il calcolo di A_z è necessario prima di tutto effettuare un tipo di compito (ad esempio un compito di valutazione o *rating task*) che permetta di avere come dati finali una serie di punti nello spazio ROC, corrispondenti a coppie di probabilità di hit e falsi allarmi al variare del criterio di decisione. Queste coordinate vengono quindi trasformate in punti z ed inserite nello spazio ROC binormale. A questo punto è necessario calcolare l'inclinazione e l'intercetta

della retta che meglio si adatta ai punti trovati, per poi procedere al calcolo di A_z (Tabella 2.3). Il valore di A_z varia tra 0.5, quando la ROC corrisponde alla diagonale positiva del grafico e la performance è casuale, e 1 quando la ROC si trova nell'angolo in alto a sinistra del grafico e la discriminazione è perfetta.

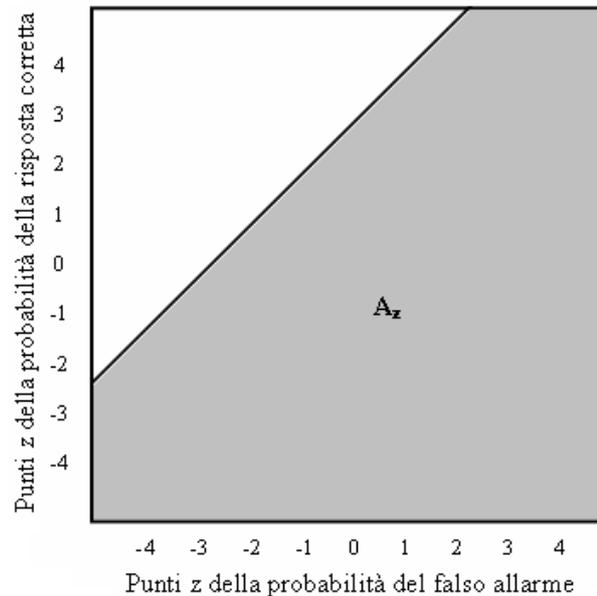


Figura 2.7. Porzione di area sotto la curva ROC in un grafico binormale.

Per il calcolo di A_z esistono una serie di programmi in cui è sufficiente inserire in *input* le frequenze delle risposte per ogni criterio utilizzato, e che forniscono in *output* l'adattamento della retta, parametri quali inclinazione e intercetta, ed una stima di A_z e della sua varianza (Stanislaw & Todorov, 1999).

L'indice A_z è associato ad un modello che non parte dall'assunto della normalità delle distribuzioni, ma piuttosto da distribuzioni di qualsiasi forma purché possano essere trasformate monotonicamente in una distribuzione normale. Laddove l'indice A_z ha certamente il vantaggio di essere indipendente dalle distribuzioni sottostanti, in realtà non risolve il problema connesso al calcolo, poiché, per poter essere calcolato, necessita che si abbia a disposizione l'intera curva ROC e non una sola coppia di valori di hit e falsi allarmi.

Swets (1996) riporta una serie di studi in cui è stato calcolato l'indice A_z e discute il fatto che i valori ottenuti nelle ricerche empiriche mostrano dei limiti: non risultano perfettamente attendibili, nel senso di poter essere ripetuti in varie prove effettuate dallo stesso sistema diagnostico, e neppure perfettamente validi, nel senso di misurare esattamente quello che si suppone misurino. Secondo l'autore però queste difficoltà non sono tanto a carico della misura utilizzata, quanto nella qualità dei dati che vengono raccolti.

In questo senso Swets (1996) propone almeno quattro possibili accorgimenti atti a migliorare la raccolta dei dati in questo tipo di rilevazioni: 1) Un problema è dato dall'adeguatezza della verità: l'osservatore dovrebbe sapere con certezza se ciascun item proposto è positivo o negativo; 2) La verità riguardante gli item dovrebbe essere determinata in modo indipendente dalle decisioni prese dall'osservatore; 3) Le procedure utilizzate per determinare la verità non dovrebbero influire sulla selezione dei casi, altrimenti si corre il rischio che la necessità di avere una verità adeguata possa inficiare il campione di item selezionato; 4) Il campione di item dovrebbe essere rappresentativo della popolazione di casi a cui il sistema diagnostico si rivolgerebbe normalmente, in altre parole i vari tipi di eventi dovrebbero aver luogo nel campione con la stessa proporzione con cui si presentano nel mondo reale.

2.3.2 Misure del criterio di decisione

Uno dei punti di partenza del presente lavoro è che la conoscenza dell'accuratezza della discriminazione non sia un requisito sufficiente a predire quale sarà il comportamento dell'osservatore. Per conoscere adeguatamente il processo di decisione è necessario avere a disposizione anche una stima del criterio che l'osservatore utilizza.

I vari indici messi a punto per ottenere una misura del criterio di decisione possono essere, prima di tutto, classificati in base a due caratteristiche fondamentali (Tabella 2.4):

Tipo di misura, che può essere: la posizione del criterio, la posizione del criterio in rapporto all'accuratezza ed il rapporto di verosimiglianza.

Forma delle distribuzioni sottese: gaussiana o normale, logistica e rettangolare.

Per quanto concerne gli indici della prima categoria, in termini generali, la posizione del criterio viene misurata come distanza tra l'asse relativa allo stesso e il punto di incrocio delle due distribuzioni.

L'indice c (Ingham, 1970) è una misura del criterio di decisione basata sulla posizione che sottende una distribuzione normale e corrisponde alla media delle quote di hit e falsi allarmi trasformate in punti z (Tabella 2.4). Il segno meno davanti alla formula serve per essere certi che si ottengano valori positivi di c solo quando il criterio è posizionato al di sopra del punto di incrocio delle distribuzioni.

Come anche le misure di accuratezza presentate in precedenza, tutti gli indici di criterio che sottendono distribuzioni logistiche, e che verranno discussi in questo paragrafo, si rifanno alla Teoria della Scelta di D. Luce (1963).

Una misura della posizione del criterio che sottende distribuzioni logistiche è b , che nella Tabella 2.4 viene presentata sia in termini logaritmici che non.

Gli indici di posizione del criterio che si basano su distribuzioni rettangolari derivano tutti dalla Teoria di Doppia Soglia (*Double Threshold Theory*) e sono discussi da Macmillan e Creelman (1990).

L'indice basato su distribuzioni rettangolari che indica la posizione del criterio è k (Tabella 2.4). Trattandosi di una trasformazione lineare della proporzione di risposte positive, viene definito anche *yes rate*. Così come l'indice c della TDS gaussiana e il b della Teoria delle Scelta, anche questa misura riflette la posizione del criterio relativamente al punto di mezzo tra le due distribuzioni.

Tutti gli indici utilizzati per determinare il criterio di decisione e basati sulla posizione dello stesso, possono essere corretti per l'accuratezza

semplicemente facendo il rapporto tra l'indice di posizione e l'indice di accuratezza previsto dal modello o dalla forma delle distribuzioni sottostanti, dando così origine ad una nuova categoria di misure.

Tabella 2.4
Classificazione delle misure del criterio di decisione della TDS

Misure del criterio di decisione	Forma delle distribuzioni sottostanti		
	<u>Gaussiana</u>	<u>Logistica</u>	<u> Rettangolare</u>
<u>Indici basati sulla posizione del criterio</u>	$c = -\frac{z(H) + z(F)}{2}$ (Ingham, 1970)	$\log(b) = -\frac{1}{2} \left[\log \frac{H}{(1-H)} + \log \frac{F}{(1-F)} \right]$ $b^2 = \frac{HF}{(1-H)(1-F)}$ (Luce, 1963)	$k = \frac{1-(F+H)}{2}$ (yes rate) (Macmillan & Creelman, 1990)
<u>Indici basati sul criterio relativamente alla accuratezza</u>	$c' = \frac{c}{d'} = -\frac{1}{2} \frac{z(H) + z(F)}{z(H) - z(F)}$	$b' = \frac{1}{2} \frac{\log(b)}{\log(\alpha)}$	$k' = \frac{1}{1 + \left(\frac{F}{1-H} \right)}$ (error ratio)
<u>Indici basati sul rapporto di Verosimiglianza</u>	$\beta = \frac{\exp \left[-\frac{z^2(H)}{2} \right]}{\exp \left[-\frac{z^2(F)}{2} \right]}$	$\beta_L = \frac{H(1-H)}{F(1-F)}$ <hr/> $B_H'' = 1 - \frac{F(1-F)}{H(1-H)} \quad \text{per } H \leq (1-F)$ $B_H'' = \frac{H(1-H)}{F(1-F)} - 1 \quad \text{per } H \geq (1-F)$ (Hodos, 1970)	
		$B'' = \frac{H(1-H) - F(1-F)}{H(1-H) + F(1-F)} \quad \text{per } H \geq F$ $B'' = \frac{F(1-F) - H(1-H)}{H(1-H) + F(1-F)} \quad \text{per } H \leq F$ (Grier, 1971)	

Come mostrato dalla Tabella 2.4, l'indice c viene semplicemente suddiviso per d' , cioè la misura di accuratezza per eccellenza in un modello di TDS gaussiano. Per normalizzare l'indice b , basato sulla Teoria della Scelta (Luce,

1963) e quindi su distribuzioni logistiche, è sufficiente utilizzare la misura di accuratezza corrispondente, cioè $\log(\alpha)$. Infine l'indice di posizione del criterio della Teoria di Doppia Soglia, k , viene corretto in base a $P(C)$, cioè la misura di accuratezza per questo modello avente distribuzioni rettangolari: la misura risultante dipende solo dalla quota di errore e viene infatti definita *error ratio*.

Per quanto concerne, infine, gli indici basati sul rapporto di verosimiglianza, l'indice relativo per un modello TDS con distribuzioni normali è già stato discusso in precedenza ed è dato dal rapporto tra le due ordinate, cioè β . L'indice equivalente per distribuzioni logistiche, prende il nome di β_L (Tabella 2.4).

Nella categoria delle misure basate sul rapporto di verosimiglianza che sottendono distribuzioni logistiche, la Tabella 2.4 riporta due ulteriori indici: B'_H e B'' . Questi indici sono stati creati e utilizzati in letteratura come misure non parametriche, cioè indipendenti dalla forma delle distribuzioni sottostanti, ma si ripropone il problema dell'adeguatezza della definizione di non parametricità.

Swets in primis (1986), Macmillan e Creelman (1990, 1996) in seguito, hanno messo in discussione la non parametricità di queste misure ed hanno dimostrato che sono trasformazioni monotoniche (quindi matematicamente equivalenti) di parametri noti delle distribuzioni logistiche. Coerentemente con questa dimostrazione, questi indici sono stati considerati alla stregua di parametri che sottendono distribuzioni logistiche, nonostante non derivino dalla Teoria della Scelta (Luce, 1963).

La prima misura del criterio di decisione basata sui triangoli è l'indice B'_H di Hodos (1970), subito dopo Grier (1971) ha proposto una misura equivalente, denominata B'' , che però ha riscosso molto più successo. Questa misura varia tra -1 e +1 e per questi due estremi le risposte sono rispettivamente sempre sì e sempre no. Sia B'_H che B'' sono trasformazioni monotoniche del parametro del

modello logistico basato sul rapporto di verosimiglianza: β_L (Macmillan & Creelman, 1990 e 1996).

2.3.3 Criterio ottimale di decisione

Nel paragrafo precedente sono stati descritti indici che conducono ad una misura del criterio a posteriori, calcolata basandosi sui dati raccolti, mentre in questo paragrafo viene introdotta una misura del criterio che può essere stabilita a priori e prende il nome di criterio ottimale di decisione. Non solo, questo indice permette anche di ottenere dall'osservatore una valutazione indipendente di due fattori fondamentali nel processo di decisione: la probabilità a priori degli eventi e la stima di benefici e costi connessi rispettivamente a risposte giuste ed errate (Swets, 1996).

Per comprendere al meglio l'impatto di queste due componenti nella decisione, è sufficiente rifarsi ad un esempio classico della TDS: un osservatore radar deve decidere se la traccia luminosa apparsa sullo schermo sia rumore di fondo, oppure un velivolo potenzialmente nemico.

Supponiamo che questi stia operando in guerra, che quindi lo scopo principale sia evitare qualsiasi attacco alla propria base. In queste condizioni, l'eventuale errore che ha costi più elevati è quello di non rilevare un pericolo esistente, mentre commettere un falso allarme avrà costi modesti. Analogamente, la segnalazione corretta dell'evento target sarà la risposta corretta che porta maggiori benefici, laddove invece rifiutare correttamente l'evento non avrà gran rilevanza. In un caso di questo genere l'osservatore tenderà a rispondere prevalentemente segnalando l'evento "pericolo".

Del tutto opposto è il caso dell'operatore radar che si trovi in tempi di pace: in questa situazione rischiare di attaccare inutilmente un improbabile nemico, sprestando munizioni e forze, sarà l'opzione meno desiderabile. I costi non sono tanto a carico delle mancate rilevazioni di eventi esistenti, quanto piuttosto dei

falsi allarmi: il comportamento dell'osservatore sarà improntato a grande cautela e raramente tenderà a segnalare l'evento "pericolo".

All'interno di un qualsiasi sistema diagnostico, a parità di accuratezza, si possono avere comportamenti molto diversi e, ovviamente, anche le caratteristiche dell'osservatore, non solo quelle situazionali, possono giocare un ruolo importante.

In un compito di diagnosi l'evidenza è una questione di grado e non una scelta assoluta. Gli esiti sono situati lungo un continuum, o comunque fanno parte di più categorie. La diagnosi è un processo probabilistico e le decisioni diagnostiche vengono prese con maggiore e minore fiducia nella loro esattezza. Questo significa altresì che prendere una decisione in un senso o nell'altro in modo sistematico richiede l'adozione di una soglia lungo la scala dell'evidenza.

In un problema di scelta tra due alternative (Tabella 2.5), possono verificarsi due tipi di decisioni corrette (hit e rifiuto corretto) e due tipi di errori (falso allarme e omissione): il criterio ottimale deve produrre un bilanciamento tra le probabilità di occorrenza di questi quattro possibili esiti. Questo però non basta: il criterio può essere definito ottimale solo se considera, non solo la probabilità dell'evento che interessa, ma anche i benefici ed i costi rispettivamente delle scelte corrette ed errate (Swets, 1996).

Tabella 2.5
Matrice di decisione statistica

	Evento	
Scelta	Positivo	Negativo
Positivo	Hit	Falso allarme
Negativo	Omissione	Rifiuto corretto

Assumiamo, per capire meglio, che la variabile su cui è richiesto di prendere una decisione sia misurata lungo un continuum e che l'evento target sia chiamato positivo, mentre l'alternativo sia denominato negativo. Poniamo,

quindi, che valori alti della variabile indichino una elevata fiducia nell'evenienza dell'evento positivo e che valori bassi della stessa indichino una bassa fiducia nel verificarsi dell'evento desiderato. Rimane, comunque, la condizione di incertezza, dal momento che gli eventi positivo e negativo non risultano completamente separati e quindi alcuni valori della variabile possono in realtà derivare sia da una evenienza che dall'altra.

Nella scelta tra i due eventi entra quindi in gioco il criterio ottimale utilizzato dall'osservatore e questo criterio determinerà una precisa compensazione nelle probabilità dei quattro possibili esiti. Se il criterio ottimale di decisione utilizzato è indulgente (o *gambler*), le proporzioni di hit e di falsi allarmi saranno abbastanza alte, mentre le proporzioni di omissioni e rifiuti corretti saranno piuttosto basse. Come è ovvio, avviene esattamente il contrario per un criterio restrittivo o conservativo.

Il criterio ottimale può essere posizionato ovunque lungo il continuum dei valori della variabile, ma storicamente esistono tre diversi approcci alla decisione che determinano tre tipi diversi di criterio e possono aiutare a definire meglio l'influenza di questo parametro sulla scelta:

- Approccio di controllo di qualità nell'industria: adotta un criterio ottimale di tipo conservativo. Questo approccio riguarda il controllo della funzionalità dei macchinari utilizzati nella produzione e per risposta positiva qui si intende che il prodotto risulta difettoso. Lo scopo è quello di limitare al massimo i falsi positivi, o falsi allarmi, cioè evitare di bloccare la produzione ogni qualvolta ci sia un minimo dubbio che il prodotto non sia adeguato.
- Approccio di salvezza dal fallimento (*fail-safe*) in ingegneria: adotta un criterio ottimale *gambler*. In questo caso il controllo riguarda la funzionalità di macchinari che verificano il giusto di funzionamento di apparecchi complessi e delicati (come ad esempio gli aeroplani) ed anche in questo caso la risposta positiva denota un malfunzionamento.

L'obiettivo è quello di evitare il più possibile le omissioni, e quindi ricevere e verificare tutte le segnalazioni di un possibile problema, dal momento che i costi di una mancanza in questo frangente sono chiaramente enormi.

- Approccio simmetrico: adotta un criterio intermedio che tende a produrre la stessa quantità di errori e risposte corrette di entrambi i tipi. E' adeguato quando non vi sia una polarizzazione rispetto a costi e benefici connessi agli esiti o rispetto alle probabilità a priori. Ad esempio i dati di alcune ricerche sul test di screening per l'HIV suggeriscono che tendere verso questo tipo di criterio massimizza la percentuale di decisioni corrette (Swets, 1996).

Fino ad ora abbiamo considerato l'utilizzo della ROC per valutare l'indice di accuratezza nella discriminazione (che varia lungo la diagonale negativa), ma questo grafico è un metodo agevole per rappresentare anche la posizione di ogni possibile tipo di criterio di decisione e del criterio ottimale (che variano lungo la diagonale positiva, perpendicolarmente al primo).

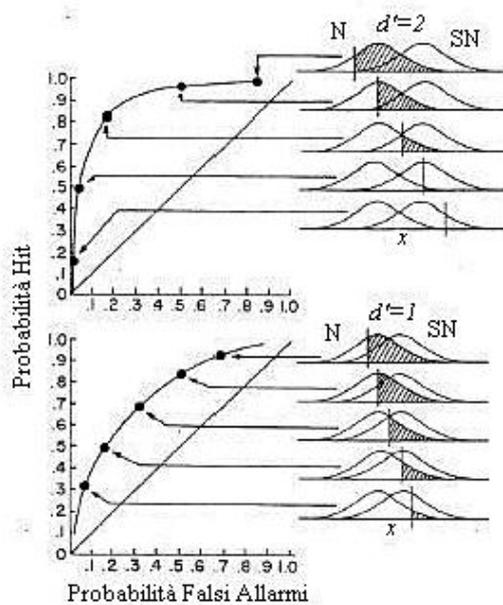


Figura 2.8. Due esempi di variazione del criterio di decisione per d' costante.

Dato che l'inclinazione della curva ROC decresce lentamente andando dall'angolo in basso a sinistra verso quello in alto a destra, il valore di inclinazione di ciascun punto della curva è un indice che descrive e rappresenta esattamente quell'unico punto. Swets (1996) definisce l'indice di inclinazione come S (da *slope*) e ritiene che esso sia una misura del criterio adottato nella scelta.

In realtà comunque la definizione del criterio ottimale prende in considerazione non solo entrambe le probabilità dei due possibili eventi, ma anche costi e benefici delle decisioni errate e corrette. Questo criterio che Swets (1996) definisce con S_{opt} , ma che noi chiameremo beta ottimale (β_{ott}), è dato dalla formula:

$$S_{opt} = \beta_{ott} = \frac{P(neg)}{P(pos)} \times \frac{(B_{tn} + C_{fp})}{(B_{tp} + C_{fn})} \quad (2.3)$$

L'alternativa negativa del processo diagnostico figura al numeratore sia come probabilità che come benefici e costi associati, mentre l'alternativa positiva è al denominatore, di nuovo, sia come probabilità che come computo costi-benefici. Da questo consegue che, a parità di costi e benefici, il valore di β_{ott} sarà alto ed il criterio conservativo quanto più la probabilità dell'evento negativo sarà maggiore di quella dell'evento positivo. Al contrario se le due probabilità vanno a favore dell'evento positivo, il valore di β_{ott} sarà basso ed il criterio *gambler*. Lo stesso può dirsi, a parità di probabilità, per le parti della formula relative a costi e benefici: il valore di β_{ott} sarà alto e il criterio conservativo quando sarà ritenuto molto più importante diagnosticare con certezza l'occorrenza dell'evento negativo. Al contrario, β_{ott} sarà basso e il criterio *gambler* quando si riterrà molto più importante essere corretti nella diagnosi dell'evento positivo.

Come esemplificato nella Figura 2.9, in generale se $\beta_{ott} < 1$ il criterio di decisione adottato può essere definito *gambler*: questo significa che si ritiene più

probabile l'evento positivo e, di conseguenza, le risposte positive (hit e falsi allarmi) saranno molto più numerose di quelle negative (rifiuti corretti e omissioni); mentre se $\beta_{ott} > 1$ il criterio di decisione è conservativo, l'evento negativo è considerato più probabile e l'osservatore deciderà più facilmente in favore di questo (alti rifiuti corretti e omissioni) che dell'evento positivo (bassi hit e falsi allarmi).

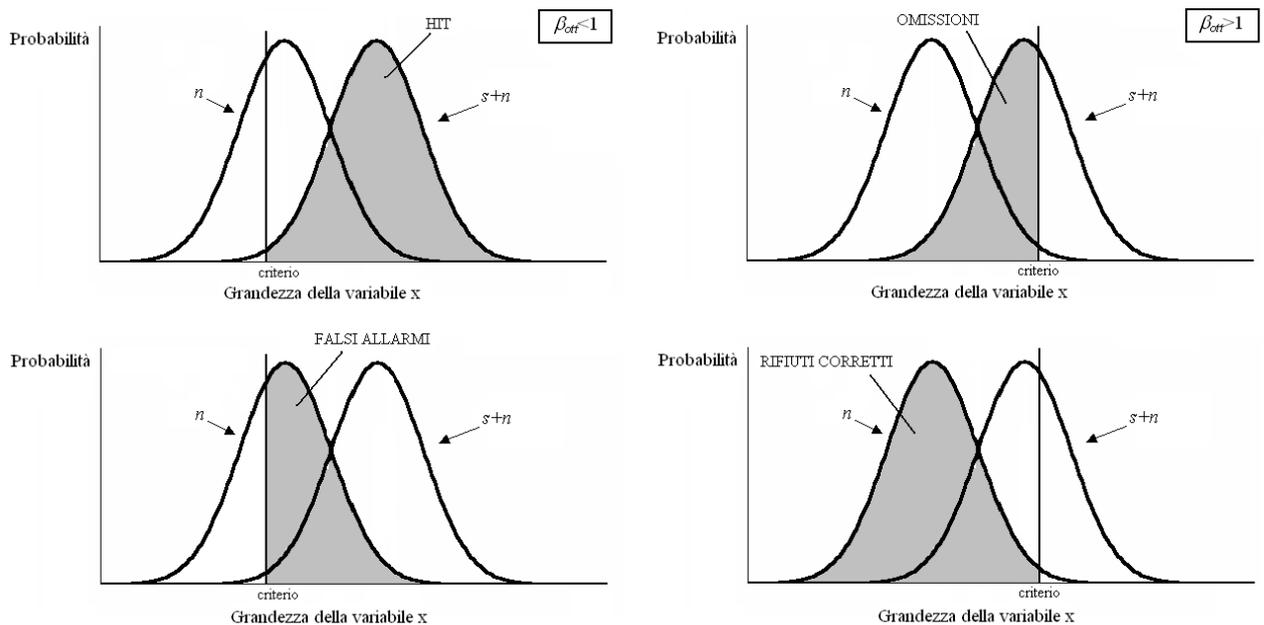


Figura 2.9. Probabilità dei quattro possibili esiti in base alla posizione del criterio ottimale di decisione.

Le stime delle probabilità e dei costi e benefici necessarie per giungere al calcolo del criterio ottimale possono ovviamente risultare piuttosto difficoltose e questo è certamente un limite da tenere in considerazione.

Nonostante questo, una posizione forte assunta da alcuni teorici della decisione, appoggiata dallo stesso Swets (1996), è che ogni scelta di un criterio di decisione implica necessariamente una serie di assunzioni in merito a probabilità, costi e benefici e che quindi rendere queste considerazioni esplicite è comunque preferibile che lasciarle tacite.

L'ottimizzazione del criterio di decisione può essere effettuata ad esempio allenando attentamente gli osservatori, fornendo loro un feedback immediato in

merito agli esiti, oppure separando i due processi di discriminazione e di decisione. Spesso accade che tutte le informazioni rilevanti per ottimizzare il criterio di decisione sono maggiormente conosciute da un decisore esterno all'osservazione che dall'osservatore stesso. Una efficiente suddivisione del compito potrebbe essere fatta fornendo una valutazione dell'osservatore in merito al valore della variabile oggetto ad un decisore che, attraverso una comparazione con il criterio di decisione, compia la scelta (Swets, 1996).

Il concetto di criterio di scelta ottimale non ha fatto presa nei numerosi e possibili ambiti diagnostici in cui potrebbe altresì risultare utile. In psicologia l'unico ambito in cui ha trovato una certa applicazione è stato nella valutazione delle attitudini per la selezione del personale; mentre per il resto se ne trovano utili applicazioni in meteorologia. Nonostante questa resistenza, il concetto non dovrebbe risultare di difficile acquisizione, molti individui ricorrono ingenuamente a questa strategia decisionale: agire in base a probabilità, costi e benefici è un modo di procedere che tutti noi utilizziamo ogni giorno. Swets (1996) stesso, per quanto consapevole delle difficoltà di riuscire ad ottenere una stima adeguata delle probabilità a priori e di costi e benefici di una serie di alternative, auspica che queste variabili non siano trattate solo in modo implicito, se non addirittura ignorate, ancora a lungo e che tutti gli strumenti a disposizione per ottenerle vengano messi a punto, studiati e perfezionati quanto prima.

Capitolo 3

La ricerca

3.1 Introduzione

Come abbiamo visto nel Capitolo 1, il rischio è stato concepito e studiato attraverso numerosi approcci, anche molto diversi tra loro. L'approccio che questa ricerca ha seguito nello studio del rischio è di tipo sperimentale: l'interesse è rivolto allo studio del comportamento e le decisioni vengono rappresentate in funzione di due componenti: la probabilità degli eventi e la valutazione costi-benefici connessa ai possibili esiti.

Queste considerazioni sono, peraltro, i punti di partenza delle Teorie dell'utilità attesa: l'individuo si comporta razionalmente e tenta di massimizzare i benefici ed operare la scelta migliore in assoluto. Per determinare la scelta ottimale, esso valuta ciascuna opzione di risposta in base sia alla probabilità che all'utilità attesa.

Allo stesso modo, anche il criterio ottimale di decisione (beta ottimale, β_{ott}), mutuato della Teoria della Detezione del Segnale (TDS), si basa su una definizione del rischio che considera contemporaneamente sia la probabilità dell'esito, che la sua desiderabilità in termini di costi e benefici. Del resto, sia la TDS che le Teorie dell'utilità attesa, rappresentano lo stato del mondo e le opzioni di risposta attraverso una tavola di decisione 2×2 .

Una critica mossa ai modelli classici di utilità attesa è quella di non tenere conto delle differenze individuali, mentre il criterio ottimale di decisione presenta la possibilità aggiuntiva di classificare il decisore in almeno due categorie: coloro che decidono di rischiare a tutti i costi (atteggiamento *gambler*) e coloro che decidono di avere una modalità improntata alla prudenza e non rischiare, se non in caso di effettivo bisogno (atteggiamento conservativo). In un certo senso queste componenti sono simili agli aspetti motivazionali che

Atkinson (1957) ha introdotto nel modello dei Valori Attesi: motivazione al successo e motivazione all'evitamento del fallimento.

Considerate le similitudini tra l'indice β_{ott} , mutuato dalla TDS, e l'approccio sperimentale, la presente ricerca è centrata su una definizione del rischio basata, per lo più, sugli aspetti comportamentali dello stesso, e che quindi non si pone l'obiettivo di indagare le componenti cognitive e motivazionali che inducono la persona a prendere una determinata decisione. Questo indice permette di misurare l'atteggiamento dell'individuo di fronte ad una scelta rischiosa, laddove non è però possibile, basandosi soltanto su questo tipo di misurazione, determinare in modo diretto la spiegazione dello stesso.

Swets (1996) ritiene che il merito di questo tipo di misurazione è quello di rendere esplicita la stima relativa alla probabilità degli eventi ed alla valutazione di costi e benefici connessi agli esiti. Partendo dall'assunto, trasversalmente condiviso, che queste componenti sono comunque alla base della decisione presa in condizioni di incertezza, risulta fondamentale sia poterle comporre in una misura unica, sia poterne idealmente scomporre gli effetti e le singole influenze nell'ambito del processo decisionale.

L'adeguatezza dell'indice β_{ott} nella misurazione del rischio è stata verificata mediante il suo utilizzo in due ambiti molto diversi tra loro, ma ugualmente importanti nello studio della presa di decisione: la professione a rischio, in particolare la professione medica, e l'adolescenza.

La scelta di due ambiti di studio così diversi è stata determinata dall'esigenza precisa di testare questo indice in vari contesti, per comprendere quanto potesse ritenersi una misura sia affidabile, nel senso di risultare ugualmente comprensibile e adatto a varie tipologie di soggetti, sia sensibile, nel senso di essere in grado di discriminare tra processi decisionali che sottendono valutazioni molto diverse tra loro.

3.2 Presa di decisione in ambito medico

Nella pratica medica prendere una decisione significa utilizzare e combinare informazioni diverse, sapere usare quelle che derivano dalle analisi mediche e dai dati clinici integrandole con la propria esperienza, riuscire a rispondere prontamente di fronte alle emergenze e selezionare le strategie decisionali più efficaci in relazione al paziente.

La presa di decisione in ambito medico ha senza dubbio le caratteristiche della decisione presa in condizioni di incertezza per la mancanza di tutte le informazioni necessarie (anche per problemi legati al tempo e ai costi). Date tali caratteristiche, il medico che compie una diagnosi, opera in uno spazio di decisione a discriminazione imperfetta. Risulta, quindi, di estrema importanza la valutazione che, in tali condizioni di incertezza, il medico compie dei costi e benefici della diagnosi (Chapman & Sonnenberg, 2000).

La Teoria della Detezione del Segnale (TDS) è sembrata particolarmente adeguata per studiare approfonditamente il criterio di decisione adottato dai medici. Peraltro questa teoria è già stata ampiamente utilizzata negli studi sulla presa di decisione nella diagnostica medica, anche se sono stati per lo più utilizzati l'analisi ROC e gli indici di accuratezza della discriminazione.

Il medico, nella sua abituale situazione di diagnosi, deve discriminare tra la presenza (M) o l'assenza (N) di una determinata malattia. A disposizione del medico sono una serie di elementi, ognuno rappresentativo di una dimensione in uno spazio multidimensionale: sintomi rilevabili all'esame obiettivo, esami di laboratorio, anamnesi, ecc.

In una situazione di questo tipo gli esiti possono essere collocati in una matrice di decisione 2x2 (Tabella 3.1), in cui le righe rappresentano la decisione del medico e le colonne lo stato del mondo.

La matrice di decisione indica quattro possibili risultati, il medico può: 1) diagnosticare correttamente la malattia (hit); 2) diagnosticare erroneamente la malattia (falso allarme); 3) non diagnosticare la malattia sbagliando perchè in

realtà il problema esiste (omissione); 4) ritenere correttamente che la malattia non vi sia (rifiuto corretto).

Tabella 3.1
Matrice di decisione statistica:
Diagnosi (malattia – assenza di malattia) x Eventi (malattia – assenza di malattia)

		<u>Eventi</u>	
		M (malattia)	N (assenza di malattia)
<u>Diagnosi</u>	M (malattia)	Hit (Diagnosi corretta)	Falso Allarme (Diagnosi in assenza di malattia)
	N (assenza di malattia)	Omissione (No diagnosi in presenza di malattia)	Rifiuto Corretto (No diagnosi in assenza di malattia)

E' possibile inoltre rappresentare la probabilità dei due eventi M e N come due distribuzioni (Figura 3.1).

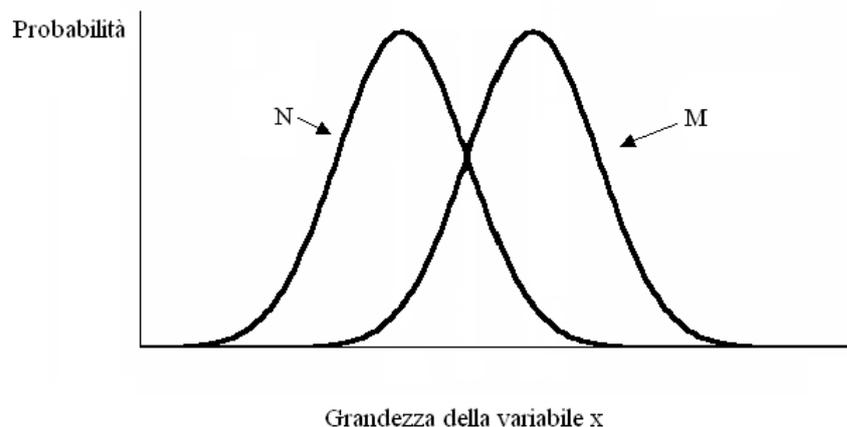


Figura 3.1 Distribuzioni delle funzioni di densità di probabilità degli eventi malattia (M) e assenza di malattia (N).

Le due distribuzioni sono in parte sovrapposte e tale sovrapposizione è tanto maggiore quanto meno il soggetto è in grado di discriminare tra le due

classi di eventi. Come abbiamo visto, esistono numerose misure, più o meno efficaci, dell'accuratezza della discriminazione, ma l'assunto da cui parte questo lavoro è che la conoscenza del grado di accuratezza non è sufficiente a predire quale sarà il comportamento del medico: è necessario disporre di una misura del criterio ottimale che utilizza per prendere la decisione.

Specificamente, il professionista valuta le conseguenze della propria risposta in base alla situazione: molto diverso sarà l'impatto della diagnosi in uno scenario in cui il medico deve gestire un'epidemia di peste bubbonica che si sta diffondendo, piuttosto che una malattia non mortale in una situazione di scarsità di mezzi e farmaci. Nel primo caso una diagnosi rapida è senza dubbio necessaria, nel secondo probabilmente è più grave una diagnosi errata che comporti uno spreco di farmaci.

L'obiettivo di questa ricerca è determinare il criterio ottimale di decisione utilizzato dai medici in fase diagnostica, ovvero comprendere dove ed in base a quale processo il medico pone il parametro beta ottimale (β_{ott}), descritto in precedenza ed utilizzato dalla TDS, appunto, come misura del criterio ottimale di decisione. Questo indice, infatti, dovrebbe permettere di valutare le considerazioni relative a costi e benefici che il medico presumibilmente effettua nel momento in cui deve prendere una decisione diagnostica. In questo studio, con i termini di "costi" e "benefici" facciamo riferimento all'accezione rispettivamente di effetti dannosi oppure positivi che incidono sulla qualità di vita del paziente.

Nell'ottica dell'utilizzo del parametro β_{ott} come indice adatto a valutare l'atteggiamento del medico nell'assunzione di rischio durante la presa di decisione, è innanzitutto necessario verificare la bontà di tale indice in questo ambito. Riuscire a proporre un utilizzo della TDS in questo senso risulta, infatti, uno degli obiettivi primari del presente studio.

L'indice β_{ott} assume in questo ambito un significato peculiare: esprime la volontà del medico di fare una diagnosi a tutti i costi anche rischiando di

sbagliare (criterio *gambler*) oppure di assumersi il rischio di non fare una diagnosi anche quando potrebbe rivelarsi necessario (criterio conservativo). Partendo dall'assunto che il criterio decisionale del medico si sarebbe distribuito lungo questo continuum, si è voluto verificare cosa accade quando le patologie hanno un diverso grado di gravità. E' stato inoltre indagato il peso dell'esperienza lavorativa nella percezione del rapporto costi-benefici nella decisione diagnostica.

Riguardo a questo aspetto, gli studi condotti (Dawson, 2000; Dawson, Connors, Speroff, Kemka, Shaw & Arkes, 1993; Berwick & Thibodeau 1983) sembrano dimostrare che il livello generale di esperienza di un medico non abbia una forte influenza sul grado di accuratezza del giudizio riguardo la prognosi. Piuttosto, sembra essere la presenza ed il grado di esperienze specifiche, ad esempio l'esperienza precedente con i pazienti deceduti, ad influire in modo marcato sull'accuratezza nella prognosi. Per quanto concerne invece l'accuratezza della diagnosi, il livello generale di esperienza potrebbe giocare un ruolo importante, ma i risultati sono controversi e necessitano di ulteriori verifiche.

Un ulteriore obiettivo del presente lavoro è quindi quello di verificare l'influenza del livello generale di esperienza nella percezione dei costi e benefici al momento della diagnosi. In questa ottica sono stati condotti due studi: uno rivolto ai medici, l'altro rivolto a studenti di medicina, che ancora non dispongono della pratica lavorativa.

3.2.1 Studio 1

Obiettivi

L'obiettivo è stato quello di valutare il criterio adottato dal medico nella formulazione della diagnosi, e verificare se questo stesso criterio varia in base alla gravità della patologia oggetto di indagine.

Metodo

Partecipanti

Hanno partecipato 42 medici (25 maschi e 17 femmine) con età media di 42.9 anni e ds di 8.4. Il tempo medio di attività svolta è 15.6 anni e ds di 8.7. All'interno del campione 31 medici sono specializzati e si distribuiscono su 13 diversi indirizzi di specializzazione (pediatria, ginecologia ed ostetricia, chirurgia, cardiologia, tossicologia, malattie dall'apparato respiratorio, allergologia, dermatologia, psichiatria, odontoiatria, neurologia, medicina estetica ed ematologia).

Strumenti

E' stato predisposto un questionario composto da due sezioni.

Nella prima è stata misurata la valutazione in termini di gravità e frequenza di un campione di 9 patologie, precedentemente selezionate (Tabella 3.2). La prova consisteva nel mettere in ordine le patologie presentate in termini di gravità ed in seguito, sempre per le stesse patologie, stabilire un ordine in base alla frequenza.

Tabella 3.2
Lista di patologie da ordinare per gravità e frequenza

<i>Qui di seguito sono elencate alcune malattie, le chiediamo, in base alla sua esperienza, di metterle in ordine in base alla gravità e in base alla frequenza.</i>
<ul style="list-style-type: none"> • AIDS • CA polmonare • TBC polmonare • Pancreatite acuta • Epatopatie • Cardiopatie • Artrite reumatoide • Diabete mellito • Influenza

Nella seconda parte veniva misurata la valutazione dei costi e benefici della diagnosi in relazione alle patologie giudicate rispettivamente di gravità lieve,

media e alta. Gli esiti, corrispondenti a ciascuna delle quattro celle della matrice di decisione della TDS, sono stati presentati sotto forma di item (Tabella 3.3). Il compito del partecipante era di valutare ciascuna affermazione su una scala da 1 a 10. Le 4 affermazioni erano ripetute per tre livelli di gravità della malattia: il medico doveva compiere la valutazione costi-benefici in relazione alla malattia da lui stesso giudicata come lieve, media e grave.

Tabella 3.3
Item relativi agli esiti della matrice di decisione 2x2:
costi delle risposte errate e benefici delle risposte corrette

<p><i>Considerando la malattia per lei più grave / meno grave / di gravità media valuti su una scala da 1 a 10:</i></p> <ul style="list-style-type: none"> • Quanto è importante una diagnosi rapida della malattia anche a costo di sbagliare con un malato che non è affetto. • Quanto è grave diagnosticare erroneamente questa malattia a un paziente che non è affetto, per cercare comunque di fare una diagnosi precoce. • Quanto è importante evitare di creare allarme non diagnosticando questa malattia e rassicurando così il paziente. • Quanto è grave non riuscire a diagnosticare questa malattia magari per un eccesso di prudenza per non allarmare il paziente.
--

Procedura

I medici sono stati contattati ed intervistati sul luogo di lavoro (ambulatorio e ospedale). La somministrazione è avvenuta in forma anonima, individuale e autosomministrata.

Analisi dei dati

Prima di presentare i risultati occorre fare alcune precisazioni sul metodo utilizzato per calcolare il β_{ott} . Come già detto, la formula del criterio ottimale di decisione è:

$$\beta_{ott} = \frac{P(neg)}{P(pos)} \times \frac{(B_m + C_{fp})}{(B_{tp} + C_{fn})} \quad (3.1)$$

Nella prima parte della formula, costituita dalle probabilità connesse alla presenza o assenza degli eventi, la frequenza attribuita da ciascun medico alla specifica patologia è stata utilizzata come stima della probabilità. Nella seconda parte della formula, invece, sono stati utilizzati i valori, da 1 a 10, che il medico aveva attribuito a ciascuna delle 4 affermazioni (Tabella 3.3) in relazione a quella specifica patologia. In questo modo è stato ottenuto un valore del criterio ottimale di decisione per ciascuna delle tre categorie di malattia: β_{ott} per le malattie gravi, β_{ott} per le malattie medie e β_{ott} per le malattie lievi.

E' stato inoltre ritenuto utile trasformare i β_{ott} , ottenuti col procedimento appena descritto, nei logaritmi naturali degli stessi ($\ln\beta_{ott}$). Quello che ha indotto a considerare opportuna questa modifica sono le caratteristiche stesse del criterio ottimale di decisione. Infatti, il β_{ott} è una distribuzione di probabilità e quindi varia tra 0 e $+\infty$, inoltre i valori si addensano per lo più attorno ad 1 creando una distribuzione asimmetrica a sinistra. La trasformazione nei logaritmi naturali consente di poter allargare la curva di distribuzione dei dati ad un *range* che varia da $-\infty$ a $+\infty$ e, allo stesso tempo, di rendere la distribuzione stessa più simile alla normale.

Laddove per il β_{ott} calcolato con la formula originale il valore soglia, che discrimina un atteggiamento *gambler* da uno conservativo, è 1 ($\beta_{ott}<1$ corrisponde ad un criterio *gambler*, mentre $\beta_{ott}>1$ ad un criterio conservativo), nel caso della trasformazione logaritmica ($\ln\beta_{ott}$) questa soglia diventa 0 (dato che $\ln 1=0$). Quindi $\ln\beta_{ott}<0$ denota un criterio *gambler*, mentre un $\ln\beta_{ott}>0$ segnala un criterio conservativo.

La trasformazione ha portato alla creazione di tre nuove variabili corrispondenti ai logaritmi naturali dei β_{ott} per i tre livelli di gravità della malattia: alta, media e bassa. Tutte le ulteriori analisi dei dati sono state effettuate utilizzando direttamente i $\ln\beta_{ott}$.

Principalmente è stata effettuata una Analisi della Varianza per misure ripetute utilizzando come variabile dipendente il valore dei $\ln\beta_{ott}$ sui tre livelli di gravità della malattia. Le analisi dei dati sono state condotte utilizzando il programma di analisi statistica SPSS 11.5 (SPSS, 2000).

Risultati

Nella Tabella 3.4 sono presentati i valori medi del $\ln\beta_{ott}$ per ciascun tipo di malattia: per le malattie gravi risulta un valore negativo, corrispondente ad un criterio *gambler*, che si attenua nelle malattie di gravità media, mentre per le malattie lievi $\ln\beta_{ott}$ diventa nettamente positivo e indica un criterio conservativo.

Tabella 3.4
Valori medi dell'indice $\ln\beta_{ott}$ in base alla gravità della malattia (medici).

	Gravità della malattia					
	Alta		Media		Bassa	
	Media	DS	Media	DS	Media	DS
$\ln\beta_{ott}$	-.82	1.29	-.11	1.41	1.98	.74

I risultati dell'Anova per misure ripetute mostrano l'esistenza di una differenza statisticamente significativa ($F_{(2,62)} = 46.49$, $p < .001$) nei valori del $\ln\beta_{ott}$ in base alla gravità della malattia. In particolare, dai confronti a coppie, emerge una differenza significativa tra malattie lievi e medie ($t_{(31)} = -7.72$, $p < .001$) e tra lievi e gravi ($t_{(33)} = -10.97$, $p < .001$).

Discussione

I risultati ottenuti in questo esperimento hanno permesso di differenziare l'atteggiamento dei medici, inteso come disponibilità all'assunzione di rischio nella diagnosi, in base al tipo di malattia, intesa, a sua volta, come gravità della malattia rispetto alla quale si trovano a dover decidere.

In particolare, nelle malattie gravi i medici scelgono un criterio ottimale di decisione *gambler*, che consiste nello stabilire la presenza di malattia a costo di commettere molti falsi allarmi. Nelle malattie lievi il criterio ottimale di

decisione diventa conservativo, preferendo non diagnosticare la malattia e commettendo, quindi, pochi falsi allarmi anche a costo di molte omissioni.

Questo diverso atteggiamento in funzione della gravità della patologia deriva probabilmente da una concezione del rapporto rischio-beneficio che il medico ha acquisito. A questo punto l'interesse è stato quello di verificare se questa sensibilità fosse legata all'esperienza lavorativa, ovvero al feedback che il medico riceve quotidianamente sugli effetti e le conseguenze delle proprie decisioni, più che alle conoscenze mediche, intese in senso di nozioni teoriche acquisite. A questo scopo è stato messo a punto ed effettuato un secondo esperimento su un campione di studenti di medicina, che si ritiene abbiano acquisito almeno in parte le nozioni teoriche, senza aver maturato un'esperienza pratica.

3.2.2 Studio 2

Obiettivi

L'obiettivo è stato quello di valutare il criterio di decisione che gli studenti di medicina adotterebbero in fase diagnostica, e verificare se questo stesso criterio varia in base alla gravità della patologia oggetto di indagine.

Metodo

Partecipanti

Hanno partecipato 26 studenti (6 maschi e 20 femmine), con età media di 23 anni e ds di 2.3, iscritti al Corso di Laurea di Medicina delle Università di Firenze e Pisa, rispettivamente 3 al 1° anno, 3 al 2° anno, 6 al 3° anno, 3 al 4° anno, 3 al 5° anno ed infine 8 al 6°anno.

Strumenti

Il questionario somministrato agli studenti è il medesimo strumento descritto ed utilizzato nel precedente esperimento. In particolare, non si ha alcuna variazione delle due sezioni che maggiormente interessano: quella che include la valutazione in termini di gravità e frequenza di un campione di 9

patologie e quella che prevede la valutazione dei costi e benefici della diagnosi in relazione alle patologie considerate di gravità lieve, media e grave.

Procedura

Gli studenti di medicina sono stati contattati ed intervistati presso le Facoltà di Medicina delle Università di Firenze e Pisa. La somministrazione è avvenuta in forma anonima, individuale e autosomministrata.

Analisi dei dati

Sono stati calcolati i beta ottimali seguendo il medesimo procedimento descritto per lo Studio 1. Questi sono poi stati trasformati nei logaritmi naturali ($\ln \beta_{ott}$), che sono stati utilizzati per le successive analisi.

Sono state effettuate una ANOVA per misure ripetute per verificare l'influenza della gravità della malattia sul criterio di decisione ed una ANOVA mista per verificare contemporaneamente l'influenza della gravità e della esperienza lavorativa sul criterio. Le analisi dei dati sono state condotte utilizzando il programma di analisi statistica SPSS 11.5 (SPSS, 2000).

Risultati

Nella Tabella 3.5 sono presentati i valori medi del $\ln \beta_{ott}$ per ciascun tipo di malattia. Per le malattie gravi $\ln \beta_{ott}$ ha valore negativo, quindi denota un criterio di decisione *gambler*, che rimane pressoché inalterato nelle malattie di media gravità, mentre per le malattie lievi diviene positivo e corrisponde ad un criterio conservativo.

Tabella 3.5
Valori medi dell'indice $\ln \beta_{ott}$ in base alla gravità della malattia (studenti di medicina).

	Gravità della malattia					
	Alta		Media		Bassa	
	Media	DS	Media	DS	Media	DS
$\ln \beta_{ott}$	-.69	1.27	-.58	1.54	1.75	.53

I risultati dell'Anova a Misure ripetute mostrano una differenza significativa ($F_{(2,40)} = 27.87, p < .001$) nei valori del $\ln\beta_{ott}$ in base alla gravità della malattia. Dall'analisi dei confronti a coppie, emerge una differenza significativa tra malattie lievi e medie ($t_{(20)} = -6.05, p < .001$) e tra lievi e gravi ($t_{(21)} = -7.22, p < .001$).

Nella Tabella 3.6 sono presentati i valori medi dei $\ln\beta_{ott}$ per i tre livelli di gravità della malattia nei due campioni: medici e studenti. I valori medi dei $\ln\beta_{ott}$ sono sempre inferiori a 0 per patologie ritenute molto o mediamente gravi, mentre sono sempre ampiamente maggiori di 1 per patologie lievi.

Tabella 3.6
Valori medi dell'indice $\ln\beta_{ott}$ in base alla gravità della malattia (medici e studenti di medicina).

	Gravità della malattia					
	Alta		Media		Bassa	
	Media	DS	Media	DS	Media	DS
$\ln\beta_{ott}$ medici	-0.82	1.29	-0.11	1.41	1.98	.74
$\ln\beta_{ott}$ studenti	-0.69	1.27	-0.58	1.54	1.75	.53

Utilizzando un procedimento di Anova Mista (variabile dipendente per misure ripetute, variabile indipendente per gruppi indipendenti), sono stati confrontati i valori dei $\ln\beta_{ott}$ ottenuti dai medici con quelli ottenuti dagli studenti per i tre livelli di gravità della malattia: i risultati mostrano che l'effetto principale dovuto al livello professionale non è significativo ($F_{(1,51)} = 0.98, p = .33$), che l'effetto principale dovuto alla gravità è significativo ($F_{(2,102)} = 69.55, p < .001$), che infine non vi è un'interazione tra i due effetti ($F_{(2,102)} = .81, p = .45$).

In altri termini, si riscontra nuovamente una diminuzione significativa del valore medio del $\ln\beta_{ott}$ passando da patologie poco gravi a molto gravi, mentre non vi sono differenze tra medici e studenti né complessivamente, né nei tre livelli di gravità considerati separatamente.

Discussione

I risultati ottenuti sui dati degli studenti di medicina hanno un andamento del tutto analogo a quelli ottenuti sui medici: non vi sono differenze tra questi due gruppi. L'atteggiamento degli studenti, inteso nuovamente come disponibilità all'assunzione di rischio nella formulazione della diagnosi, si differenzia in base alla gravità della malattia, in modo del tutto simile a quanto già rilevato per i medici.

In particolare, nelle malattie gravi il criterio ottimale di decisione che emerge è, di nuovo, *gambler*: stabilire la presenza di malattia a costo di commettere falsi allarmi; allo stesso modo, nelle malattie lievi il criterio è conservativo preferendo stabilire la non presenza di malattia, commettendo pochi falsi allarmi e molte omissioni.

3.3 Presa di decisione in adolescenza

Lo studio della presa di decisione e della percezione del rischio si è spesso occupato proprio della popolazione giovanile, in quanto soggetta ad assumere comportamenti rischiosi e potenzialmente lesivi (Benthin, Slovic & Severson, 1992; Savadori & Rumiati, 1996).

Si può ipotizzare che l'adolescente, nel momento in cui debba decidere se adottare o meno una condotta potenzialmente pericolosa, abbia a disposizione una serie di informazioni: sulle caratteristiche attrattive e negative dello stesso, sulla probabilità che si verifichino conseguenze negative, sull'impatto che questo avrebbe sulle figure per lui importanti (Slovic, 1987). Tutte queste informazioni dovrebbero orientare la valutazione che l'adolescente compie dei costi e dei benefici connessi ad una scelta. Nonostante questo, le informazioni che l'adolescente ha a disposizione sono delle probabilità e quindi opera nell'ambito della presa di decisione in condizioni di incertezza.

Date queste premesse, si può ritenere che la TDS, ed in particolare il criterio ottimale di decisione β_{ott} , risulti adeguata per studiare la valutazione che l'adolescente compie di rischi e benefici in merito alle sue scelte.

La presente ricerca vuole fornire un contributo alla misurazione della percezione del rischio in adolescenza, utilizzando il β_{ott} come indice per la misurazione della valutazione costi-benefici. In questo ambito, il comportamento target è la disponibilità dell'adolescente a chiedere aiuto nel caso in cui l'adozione di taluni comportamenti a rischio comportasse delle conseguenze negative di varia gravità. In questo modo, il β_{ott} potrebbe essere un indice in grado di rivelare l'atteggiamento dei giovani nel chiedere aiuto e la loro valutazione dei costi e benefici.

Se il criterio di decisione selezionato risultasse conservativo ($\beta_{ott} > 1$), questo significherebbe che l'adolescente preferisce evitare a tutti i costi di chiedere aiuto, pur assumendosi il rischio di sbagliare nel caso in cui conseguenze negative abbiano luogo. Al contrario, se il criterio risultasse *gambler* ($\beta_{ott} < 1$), l'adolescente si mostrerebbe propenso a richiedere aiuto con facilità, pur rischiando magari di prevenire conseguenze che non avrebbero mai avuto luogo.

In sintesi, in questo studio la presa di decisione concerne la disponibilità o meno a chiedere aiuto di fronte a possibili conseguenze negative derivanti da comportamenti a rischio; e la valutazione costi-benefici riguarda l'opportunità della richiesta di aiuto, in vista delle conseguenze che questo tipo di autodenuncia potrebbe comportare.

La matrice di decisione statistica, che racchiude eventi e possibili risposte di un compito di questo genere, è riprodotta dalla Tabella 3.7.

L'adolescente può quindi trovarsi di fronte a quattro possibili esiti: una richiesta di aiuto corretta (hit), una richiesta di aiuto non necessaria (falso allarme), una mancanza di richiesta di aiuto 'errata', nel senso che la conseguenza negativa ha luogo e sarebbe stato quindi necessario chiedere aiuto (omissione), ed una mancanza di richiesta di aiuto 'corretta', nel senso che la conseguenza negativa non ha avuto luogo e quindi non c'è bisogno di aiuto (rifiuto corretto).

Tabella 3.7
 Matrice di decisione statistica:
 Risposte (richiesta di aiuto – no richiesta di aiuto) x
 Eventi (conseguenze negative – no conseguenze negative)

	<u>Eventi</u>	
<u>Risposte</u>	Conseguenze Negative	No Conseguenze Negative
Richiesta di aiuto	Hit (Richiesta di aiuto corretta)	Falso Allarme (Richiesta di aiuto laddove non necessaria)
No Richiesta di aiuto	Omissione (No Richiesta di aiuto laddove necessaria)	Rifiuto Corretto (No Richiesta di aiuto laddove non necessaria)

L'obiettivo di questa ricerca è stato, in prima istanza, quello di mettere a punto un questionario che fosse di facile comprensione per gli adolescenti e permettesse di valutare la loro disponibilità alla richiesta di aiuto. Di seguito l'obiettivo è stato quello di verificare se questa propensione varia in base al tipo di comportamento a rischio indagato, alla gravità delle conseguenze possibili, al genere e all'adozione o meno del comportamento. Allo stesso tempo, è stato effettuato un controllo preliminare sulla validità del criterio ottimale di decisione.

3.3.1 Studio 1

Obiettivi

Ponendo i giovani di fronte a possibili conseguenze negative delle condotte legate ad uso ed abuso di droghe sintetiche e sostanze alcoliche, è stata innanzitutto misurata la disponibilità dei partecipanti a chiedere aiuto; in seguito è stato verificato se questa propensione varia in base alla gravità delle conseguenze proposte, al comportamento indagato ed al fatto che gli adolescenti intervistati adottino o meno quel comportamento.

*Metodo*Partecipanti

Hanno partecipato allo studio 405 adolescenti (229 maschi e 176 femmine) con età media 17.7 anni e ds di 0.7. I partecipanti frequentano le classi IV e V di due Licei Scientifici delle province di Firenze e Pisa.

Strumenti

Per ciascuna delle due aree tematiche indagate (uso di sostanze stupefacenti e uso di sostanze alcoliche) è stato predisposto un questionario composto da tre sezioni.

Nella prima si è indagato sia il comportamento e gli stili di consumo dei giovani, che le loro conoscenze in merito agli effetti, alle conseguenze e agli aspetti legali connessi all'uso di questo tipo di sostanze.

Nella seconda sezione veniva presentata a ciascun partecipante una lista di 9 conseguenze di varia gravità connesse all'abuso delle sostanze indagate (Tabella 3.8); il compito richiesto era quello di posizionare le stesse conseguenze prima in ordine di gravità, poi in ordine di frequenza.

Tabella 3.8

Lista di conseguenze, connesse all'abuso di sostanze, da ordinare per gravità e frequenza

<p><i>Ordina questo elenco di possibili conseguenze negative in base alla gravità ed alla frequenza:</i></p> <ul style="list-style-type: none"> • Disturbi fisici (nausea, tremori, sudorazione, stanchezza, tics facciali) • Danni epatici • Danni cerebrali • Disturbi del sonno • Danni nel rendimento scolastico e lavorativo • Problemi con la giustizia • Disfunzioni sessuali • Disturbi psicologici (depressione dell'umore, irritabilità, stati d'ansia) • Dipendenza

Infine nella terza ed ultima sezione, per ciascuna delle tre conseguenze che il partecipante stesso aveva classificato come più grave, meno grave o mediamente grave, sono state presentate quattro affermazioni: due corrispondenti ai rischi connessi ad una decisione sbagliata, mentre le altre due corrispondenti ai benefici connessi ad una decisione giusta. Il compito del partecipante era valutare ciascuna affermazione su una scala da 1 a 10 (Tabella 3.9).

Tabella 3.9
Item relativi agli esiti della matrice di decisione 2x2:
costi delle risposte errate e benefici delle risposte corrette

<p><i>Considerando la conseguenza che hai scelto come più grave / meno grave / di media gravità ti chiediamo di valutare su una scala da 1 a 10 ciascuna delle seguenti affermazioni:</i></p> <ul style="list-style-type: none"> • Quanto è importante chiedere aiuto quando questo tipo di conseguenza potrebbe verificarsi? • Quanto è grave chiedere aiuto quando questo tipo di conseguenza potrebbe non verificarsi? • Quanto è importante non chiedere aiuto quando questo tipo di conseguenza potrebbe non verificarsi? • Quanto è grave non chiedere aiuto quando questo tipo di conseguenza potrebbe verificarsi?
--

Procedura

Il questionario era in forma anonima ed autosomministrato. La somministrazione, previo accordo con gli insegnanti, è avvenuta collettivamente in classe durante l'orario scolastico.

Analisi dei dati

Il primo passo è stato quello di ricavare dai dati ottenuti i valori dell'indice β_{ott} per ciascun partecipante e per ciascuno dei tre tipi di conseguenze prese in considerazione (gravità alta, media e lieve). Per ottenere questo indice è stata utilizzata la formula originale indicata nella TDS:

$$\beta_{ott} = \frac{P(neg)}{P(pos)} \times \frac{(B_m + C_{fp})}{(B_p + C_{fn})} \quad (3.1)$$

Come stima della probabilità dell'evento e, per sottrazione, della probabilità del non evento è stata utilizzata la frequenza attribuita da ogni partecipante al presentarsi della conseguenza prescelta; mentre nella parte destra della formula sono stati utilizzati i valori assegnati a ciascuna delle quattro affermazioni corrispondenti alle celle della matrice di decisione (Tabella 3.9).

Al termine di questa operazione di codifica, sono state ottenute tre nuove variabili: β_{ott} per le conseguenze di gravità alta, β_{ott} per le conseguenze di gravità media e β_{ott} per le conseguenze di gravità lieve.

Preliminarmente sono state condotte analisi dei dati di tipo descrittivo, volte a verificare la normalità delle distribuzioni delle variabili ed a rilevare la presenza di eventuali missing ed outliers che avrebbero potuto inficiare le analisi successive.

Sono state, quindi, effettuate Analisi della Varianza per misure ripetute sia sul campione totale, che sui due sottocampioni corrispondenti ai comportamenti indagati (uso di droghe e uso di alcool) considerati separatamente, per verificare l'influenza dei tre livelli di gravità sul criterio ottimale di decisione.

Le analisi dei dati sono state condotte utilizzando il programma di analisi statistica SPSS 11.5 (SPSS, 2000).

Risultati

Nella Tabella 3.10 sono presentati i valori medi dei β_{ott} per i tre livelli di gravità della conseguenza. In particolare, per conseguenze molto gravi o, al contrario, poco gravi i β_{ott} assumono valori superiori ad 1 e questo indica che i partecipanti tendono ad avere un atteggiamento conservativo. Nel caso di conseguenze considerate di media gravità, invece, i β_{ott} assumono valori inferiori ad 1, indicando un atteggiamento *gambler*.

Tabella 3.10
Valori medi dell'indice β_{ott} in base alla gravità della conseguenza.

	Gravità della conseguenza					
	Alta		Media		Bassa	
	Media	DS	Media	DS	Media	DS
β_{ott}	1.95	2.1	.77	.87	1.53	2.4

Dalle analisi condotte sul campione totale (ANOVA per misure ripetute) emerge una differenza statisticamente significativa ($F_{(2,602)}= 25.99, p<.001$) tra i β_{ott} di conseguenze ritenute rispettivamente molto, poco o mediamente gravi. Inoltre dai confronti a coppie emergono delle differenze statisticamente significative tra conseguenze lievi e medie ($t_{(310)}=5.21, p<.001$) e tra conseguenze medie e gravi ($t_{(307)}=9.02, p<.001$).

Risultati analoghi sono stati riscontrati per i due sottocampioni considerati separatamente (Tabella 3.11). Anche in questo caso, infatti, per conseguenze ritenute gravi e lievi il β_{ott} è sempre superiore ad 1, mentre per conseguenze di media gravità il β_{ott} è inferiore ad 1.

Tabella 3.11
Valori medi dell'indice β_{ott} in base alla gravità della conseguenza e per tipo di comportamento indagato.

	Gravità della conseguenza					
	Alta		Media		Bassa	
	Media	DS	Media	DS	Media	DS
β_{ott} campione droga	1.31	1.73	.79	.98	1.84	2.62
β_{ott} campione alcool	2.46	2.31	.75	.78	1.29	2.28

All'ANOVA per misure ripetute emergono nuovamente delle differenze statisticamente significative tra i β_{ott} di conseguenze considerate molto, poco o

mediamente gravi sia nel sottocampione droga ($F_{(2,262)}= 9.004, p<.001$) che nel sottocampione alcool ($F_{(2,338)}= 31.925, p<.001$).

Per quanto concerne i confronti a coppie, nel caso del sottocampione riferito all'utilizzo di droga i risultati sono analoghi a quelli del campione totale: vi sono delle differenze statisticamente significative tra i valori del criterio β_{ott} tra conseguenze di gravità alta e media ($t_{(135)}=2.89, p<.05$) e tra conseguenze di gravità media e bassa ($t_{(136)}=4.31, p<.001$).

Nel sottocampione riferito all'utilizzo di alcool, tutti e tre i confronti effettuati risultano statisticamente significativi: tra conseguenze lievi e medie ($t_{(173)}=3.07, p<.05$), lievi e gravi ($t_{(172)}=4.11, p<.001$), medie e gravi ($t_{(171)}=9.63, p<.001$).

E' stata quindi effettuata una ANOVA mista sul campione totale: tenendo la gravità della conseguenza come variabile dipendente per misure ripetute ed inserendo il tipo di comportamento indagato (droga o alcool) come variabile indipendente. L'effetto principale dovuto alla gravità è risultato statisticamente significativo ($F_{(2,600)}= 24.21, p<.001$), l'effetto principale dovuto al tipo di comportamento indagato non è risultato significativo ($F_{(1,300)}= 2.78, p=.096$), mentre l'interazione tra i due effetti è risultata statisticamente significativa ($F_{(2,600)}= 13.82, p<.001$). Dall'analisi di questo effetto di interazione (Tabella 3.11) emerge che: per le conseguenze considerate gravi il valore medio del β_{ott} per il sottocampione alcool è sensibilmente più alto rispetto al sottocampione droga, e questo andamento si inverte per conseguenze considerate lievi, mentre per conseguenze di media gravità il β_{ott} assume un valore sostanzialmente identico. Tuttavia un successivo confronto a coppie ha permesso di stabilire che tra i due comportamenti esiste una differenza statisticamente significativa ($t_{(315)}=4.73, p<.001$) solo per quanto riguarda le conseguenze ritenute gravi.

Infine, all'interno del sottocampione droga, è stato verificato se vi fossero delle differenze significative nella stima dei β_{ott} da parte di coloro che utilizzano abitualmente (*participants*) o meno (*non-participants*) alcune sostanze. In

particolare, sono state verificate le differenze tra consumatori e non consumatori di *cannabis* e tra consumatori e non consumatori di nicotina.

Nel caso della *cannabis*, la Tabella 3.12 riassume i valori medi dei β_{ott} sui tre livelli di gravità distinguendo tra *participants* e *non-participants*. Come si può notare, coloro che consumano abitualmente *cannabis* hanno sempre un punteggio medio più alto rispetto ai *non-participants*.

Tabella 3.12
Valori medi dell'indice β_{ott} in base alla gravità della conseguenza e per consumo o meno di *cannabis*.

		Gravità della conseguenza					
		Alta		Media		Bassa	
		Media	DS	Media	DS	Media	DS
Cannabis	β_{ott} <i>participants</i>	1.882	2.94	1.56	2.48	3.46	5.48
	β_{ott} <i>non-participants</i>	1.451	1.98	.92	1.47	1.93	4.54

E' stata quindi effettuata una ANOVA mista (gravità della conseguenza su misure ripetute x consumo su due livelli). I risultati mostrano la significatività dell'effetto principale dovuto alla gravità ($F_{(2,290)} = 6.53, p < .05$), la significatività dell'effetto principale dovuto all'utilizzo o meno della sostanza ($F_{(1,145)} = 7.36, p < .05$), la mancanza di interazione tra questi due effetti ($F_{(2,290)} = .98, p = .37$). Questi risultati, oltre a confermare l'effetto derivato dalla gravità della conseguenza, mostrano che, nel complesso, i consumatori di *cannabis* hanno un atteggiamento maggiormente reticente a richiedere aiuto rispetto ai coetanei non consumatori.

Nel caso del consumo di nicotina i valori medi del β_{ott} , suddivisi per gravità e consumo, sono presentati nella Tabella 3.13. Anche in questo caso è stata effettuata una ANOVA mista ed i risultati sono analoghi a quanto detto per la *cannabis*: l'effetto principale dovuto alla gravità è significativo ($F_{(2,290)} = 6.62$,

$p < .05$), l'effetto principale dovuto al consumo è significativo ($F_{(1,145)} = 6.74$, $p < .05$), mentre l'effetto di interazione non è statisticamente significativo ($F_{(2,290)} = 3.25$, $p = .055$). Anche in questo secondo caso i risultati confortano l'evidenza che i consumatori hanno sostanzialmente un atteggiamento più conservativo rispetto alla richiesta di aiuto che i non consumatori.

Tabella 3.13
Valori medi dell'indice β_{out} in base alla gravità della conseguenza e per consumo o meno di nicotina.

		Gravità della conseguenza					
		Alta		Media		Bassa	
		Media	DS	Media	DS	Media	DS
Nicotina	β_{out} <i>participants</i>	1.56	2.46	1.46	2.34	3.52	6.55
	β_{out} <i>non-participants</i>	1.63	2.29	.89	1.42	1.63	2.84

Discussione

La distribuzione dei valori di β_{out} mantiene un trend costante: i partecipanti hanno mostrato un atteggiamento conservativo, quindi una minor propensione a chiedere aiuto, se posti di fronte a possibili conseguenze molto dannose o poco dannose. Al contrario, i partecipanti sembrano essere maggiormente orientati a chiedere aiuto, quindi ad esporsi, nel caso in cui le conseguenze derivanti da un comportamento a rischio siano di media gravità.

Inoltre, per le conseguenze ritenute gravi, gli adolescenti sembrano assumere un atteggiamento maggiormente conservativo, riluttante alla richiesta di aiuto, quando la conseguenza negativa deriva da comportamenti di abuso di sostanze alcoliche, piuttosto che nella situazione in cui tale conseguenza derivi dall'abuso di sostanze stupefacenti.

Un altro dato interessante concerne il fatto che i consumatori di *cannabis* e di nicotina sembrano assumere un atteggiamento particolarmente conservativo

rispetto alla richiesta di aiuto, mentre questa modalità è molto meno accentuata in coloro che non utilizzano abitualmente questo tipo di sostanze.

Nonostante questo primo studio abbia ottenuto dei risultati interessanti, sono state riscontrate alcune difficoltà a livello della comprensibilità del questionario da parte degli adolescenti. Le difficoltà relative alla modalità di indagine potrebbero essere, inoltre, tra le cause che hanno determinato una ampia variabilità dei valori ottenuti dell'indice β_{ott} . Allo scopo di migliorare l'affidabilità della misurazione è stato quindi deciso di condurre un secondo studio, nel quale il questionario è stato preliminarmente modificato per incrementarne la comprensibilità. Inoltre in questo secondo studio è stato effettuato anche un controllo sulla validità del β_{ott} .

3.3.2 Studio 2

Obiettivi

Questo secondo studio ha come obiettivi: 1) la messa a punto del questionario con l'introduzione di supporti visivi, nella forma di vignette, che rendessero maggiormente comprensibili gli item relativi ai costi e benefici corrispondenti alle celle della matrice di decisione statistica; 2) la misurazione, mediante l'indice β_{ott} , della disponibilità dell'adolescente a chiedere aiuto di fronte a conseguenze negative derivanti da comportamenti a rischio; 3) l'introduzione di una misura della propensione alla ricerca di sensazioni forti in adolescenza da utilizzare come contributo alla validità dell'indice β_{ott} . La scala introdotta nel protocollo a questo scopo è la *Sensation Seeking Scale* di Zuckerman (1978).

Metodo

Partecipanti

Hanno partecipato 444 adolescenti (188 maschi e 256 femmine) di età media 18 anni e ds di 1.1. I partecipanti frequentano le classi IV e V di un

Istituto Tecnico Industriale e di un Liceo Pedagogico e Linguistico della provincia di Pisa.

Strumenti

Al fine di migliorare la comprensibilità del questionario, è stata creata una vignetta che riproduce un dialogo tra due personaggi (Figura 3.2) per ciascuna delle quattro celle della matrice di decisione statistica: nello specifico per i benefici delle risposte corrette (hit e rifiuto corretto) e per i costi delle risposte errate (falso allarme e omissione).

Il primo personaggio manifesta la preoccupazione di incorrere in un disturbo fisico a causa di un comportamento a rischio e si interroga in merito all'opportunità di chiedere aiuto; il secondo personaggio risponde a questo dubbio incoraggiando o meno verso la richiesta di aiuto e portando a supporto del consiglio fornito una valutazione di costi o benefici.



Figura 3.2. Esempio di vignetta relativa ai benefici della risposta corretta (*hit*).

Il foglio di risposta conteneva in una pagina unica le quattro vignette e le quattro affermazioni e, a ciascun partecipante, è stato chiesto di collegare ogni vignetta all'affermazione ritenuta corrispondente. Hanno partecipato alla fase di messa a punto dello strumento 76 adolescenti di età media 17.7 anni e ds di .75.

Sulla base dei risultati ottenuti nella parte preliminare dello studio, le vignette sono state introdotte nel questionario definitivo (Questionario Beta).

Nel Questionario Beta, inizialmente è stato chiesto a ciascun partecipante di stimare la percentuale di ragazzi che, consumatori abituali di una determinata sostanza, rischiassero di andare incontro ad un certo tipo di conseguenza negativa (Tabella 3.14).

Tabella 3.14
Stima della percentuale di giovani consumatori a rischio

Secondo un recente studio condotto sulla popolazione giovanile, i ragazzi che consumano abitualmente droghe sintetiche (ecstasy, anfetamine, LSD), rischiano di procurarsi dei seri problemi cerebrali.

Secondo te, che percentuale dei ragazzi, consumatori abituali di droghe sintetiche, rischia di andare incontro a problemi cerebrali? _____

A questo punto è stato presentato uno scenario in cui è stata descritta la situazione di un ragazzo di 17 anni che, avendo adottato un comportamento a rischio, non si sentiva molto bene e si interrogava sull'opportunità di richiedere o meno aiuto. Questo ulteriore supporto ha avuto lo scopo di chiarire lo scenario e far immedesimare i partecipanti nella situazione. Di seguito sono state presentate quattro vignette, corredate da quattro affermazioni, corrispondenti alle celle della matrice di decisione statistica. Il compito del partecipante era valutare ciascuna vignetta, e la corrispondente affermazione, su una scala da 1 a 10.

Oltre al Questionario Beta, è stata somministrata la *Sensation Seeking Scale* di Zuckerman (1978) che misura la tendenza a ricercare e sperimentare sensazioni forti. Questa scala è composta da 40 item con modalità di risposta dicotomica, è suddivisa in quattro sottoscale, composte ciascuna da 10 item (*Thrill & Adventure Seeking, Experience Seeking, Disinhibition e Boredom Susceptibility*), e fornisce anche un punteggio totale: a punteggio alto corrisponde una marcata tendenza alla ricerca di sensazioni e viceversa.

Il disegno della ricerca è stato costruito in modo da creare sei versioni parallele del medesimo questionario. E' stato deciso di indagare nuovamente due tipi di comportamento a rischio: uso e abuso di sostanze stupefacenti ed uso e abuso di sostanze alcoliche; e di prendere in considerazione, per ciascun comportamento, tre tipi di possibili conseguenze negative di gravità variabile: danni cerebrali (gravità alta), disturbi psicologici (gravità media) e disturbi del sonno (gravità lieve).

Procedura

Il questionario era in forma anonima ed autosomministrato. La somministrazione, previo accordo con gli insegnanti, è avvenuta collettivamente in classe durante l'orario scolastico.

Analisi dei dati

Per verificare la comprensibilità delle vignette e l'equivalenza di contenuto tra vignette ed affermazioni è stata semplicemente calcolata la frequenza e la percentuale delle risposte corrette.

In seguito alla somministrazione del Questionario Beta, il primo passo è stato quello di ricavare dai dati ottenuti i valori dell'indice β_{ott} per ciascun partecipante. Per il calcolo del criterio ottimale è stata utilizzata la formula 3.1: la percentuale attribuita dal partecipante al presentarsi della conseguenza indicata è stata usata come stima della probabilità dell'evento e, per sottrazione, della probabilità del non evento; i valori assegnati a ciascuna delle quattro vignette sono stati sostituiti al posto dei costi e benefici connessi ai quattro possibili esiti. Al termine di questo procedimento è stata ottenuta una nuova variabile corrispondente al valore dell'indice β_{ott} .

Preliminarmente sono state condotte analisi dei dati di tipo descrittivo, volte a verificare la normalità delle distribuzioni delle variabili ed a rilevare la presenza di eventuali missing ed outliers che avrebbero potuto inficiare le analisi successive.

Sono stati, quindi, effettuati t di Student per campioni indipendenti: come variabile dipendente è stato usato l'indice β_{ott} , come variabili indipendenti si è scelto il comportamento indagato (uso di droghe o uso di alcool) ed il genere. Sono state poi effettuate Analisi della Varianza (ANOVA), sia sul campione totale, che sui due sottocampioni considerati separatamente, prendendo come variabile dipendente il punteggio β_{ott} e come variabile indipendente la gravità della conseguenza considerata. Infine è stata effettuata una analisi delle correlazioni (r di Pearson) tra l'indice β_{ott} e le scale della *Sensation Seeking Scale* ed una regressione lineare semplice, utilizzando come variabile dipendente l'indice β_{ott} e come variabile indipendente il punteggio totale di *Sensation Seeking*.

Le analisi dei dati sono state condotte utilizzando il programma di analisi statistica SPSS 11.5 (SPSS, 2000).

Risultati

L'esattezza nel collegare una vignetta con l'affermazione corrispondente è risultata complessivamente soddisfacente (Tabella 3.15).

Tabella 3.15
Percentuali di accordo nel collegamento tra frasi e vignette.

	Vignetta V _{hit}	Vignetta C _{fa}	Vignetta V _{rc}	Vignetta C _{omi}
Frase V _{hit}	47 61.8%	7 9.2%	4 5.3%	20 15.8%
Frase C _{fa}	2 2.6%	42 55.3%	18 23.7%	12 15.8%
Frase V _{rc}	5 6.6%	16 21.1%	51 67.1%	2 2.6%
Frase C _{omi}	22 28.9%	11 14.5%	3 3.9%	42 55.3%

Nello specifico tutte le vignette sono state correttamente attribuite alle affermazioni corrispondenti da almeno il 55% del campione. Le vignette di più facile comprensione, e quindi attribuite correttamente con maggiore frequenza, sono state quella relativa ai benefici della hit (61.8%) e quella sui benefici del rifiuto corretto (67.1%). Probabilmente la strategia di suggerire una scelta

mostrandone i lati positivi è maggiormente familiare e comprensibile, per gli adolescenti, rispetto a quella di sfavorire una scelta mettendone in evidenza gli svantaggi.

Per quanto concerne i risultati dei dati raccolti in seguito alla somministrazione del Questionario Beta, la Tabella 3.16 riassume i valori medi dell'indice β_{ott} in base al tipo di comportamento indagato. Nello specifico, nonostante il valore medio dell'indice sia sempre inferiore ad 1, che indica un atteggiamento di propensione alla richiesta di aiuto, quando si considerano le conseguenze derivanti dall'uso di droga la media è sensibilmente più bassa, rispetto a quando si considera l'alcool, e questa differenza è risultata statisticamente significativa ($t_{(442)}=2.83, p<.01$). Questo significa che, di fronte a conseguenze potenzialmente negative dovute all'utilizzo di droga, l'atteggiamento di disponibilità a richiedere aiuto sembra essere ancor più forte.

Tabella 3.16
Valori medi dell'indice β_{ott} in base al tipo di comportamento indagato.

	Comportamento			
	Droga		Alcool	
	Media	DS	Media	DS
β_{ott}	.54	1.14	.88	1.32

La grandezza dell'effetto, calcolata mediante l'indice d di Cohen, mostra che l'effetto trovato è di entità lieve ($d=.27$), mentre l'analisi della potenza del test, effettuata a posteriori, rivela una potenza pari a .88 ed una probabilità di commettere un errore di secondo tipo pari a .12. Questo significa che il test utilizzato aveva una probabilità molto alta (88%) di mettere in evidenza un effetto senza sbagliare ed analogamente una probabilità molto bassa (12%) di negare erroneamente l'esistenza di un effetto.

La Tabella 3.17 riassume i valori medi dell'indice β_{ott} in base al genere: i maschi hanno un punteggio medio più alto delle femmine. Questa differenza è risultata statisticamente significativa ($t_{(442)}=3.45, p<.01$). In particolare, le

femmine mostrano un atteggiamento di maggiore propensione alla richiesta di aiuto rispetto ai maschi.

Tabella 3.17
Valori medi dell'indice β_{ott} in base al genere.

	Genere			
	Maschi		Femmine	
	Media	DS	Media	DS
β_{ott}	.97	1.52	.53	.96

L'indice di Cohen assume valore $d=.37$ mettendo in evidenza un effetto di entità moderata, mentre l'analisi di potenza a posteriori ha valore .98 con una corrispondente probabilità di commettere un errore di secondo tipo di .02. Quindi il test aveva una probabilità estremamente alta (98%) di rilevare correttamente un effetto.

La Tabella 3.18 presenta i punteggi medi dell'indice β_{ott} in base alla gravità della conseguenza negativa (alta, media, bassa) sia nel campione totale, che nei due sottocampioni relativi al tipo di comportamento indagato (alcohol e droga).

Tabella 3.18
Valori medi dell'indice β_{ott} in base alla gravità della conseguenza nel campione totale e per tipo di comportamento indagato.

	Gravità della conseguenza					
	Alta		Media		Bassa	
	Media	DS	Media	DS	Media	DS
β_{ott} campione totale	.62	1.13	.77	1.38	.74	1.21
β_{ott} campione alcohol	.74	1.03	1.06	1.74	.84	1.11
β_{ott} campione droga	.49	1.23	.49	.83	.65	1.31

Nel campione totale i punteggi medi dell'indice sui tre livelli di gravità sono tutti inferiori ad 1 indicando un atteggiamento *gambler* rispetto alla richiesta di aiuto, ovvero una marcata disponibilità a richiedere aiuto di fronte a possibili conseguenze negative.

Dall'ANOVA non è emersa alcuna differenza significativa in base ai tre livelli della variabile gravità ($F_{(2,441)}=.63, p=.53$).

Un andamento analogo al campione totale è mostrato in Tabella 3.18 per quanto riguarda i punteggi medi nel campione droga: il valore medio dell'indice β_{ott} è sempre inferiore ad 1 e dall'ANOVA non è emersa alcuna differenza statisticamente significativa ($F_{(2,216)}=.46, p=.63$). Per quanto concerne il campione alcool, la Tabella 3.18 mostra che nel caso di conseguenze di media gravità il valore medio di β_{ott} è maggiore ad 1, denotando un atteggiamento di reticenza alla richiesta di aiuto, mentre per conseguenze di gravità alta e bassa l'indice β_{ott} è mediamente inferiore ad 1. L'ANOVA condotta su questo campione non ha comunque rilevato alcuna differenza statisticamente significativa nei punteggi β_{ott} sui tre livelli di gravità delle conseguenze ($F_{(2,222)}=1.19, p=.31$).

Infine i punteggi ottenuti all'indice β_{ott} sono stati correlati con le quattro sottoscale e con il punteggio totale della *Sensation Seeking Scale*. Sono emerse delle correlazioni basse, ma statisticamente significative, con le tre sottoscale *Thrill & Adventure Seeking* ($r=.09, p<.05$), *Experience Seeking* ($r=.11, p<.05$) e *Disinhibition* ($r=.16, <.05$) e con il totale di scala ($r=.16, <.05$). Nonostante gli indici di correlazione siano bassi, è stata effettuata una analisi di regressione semplice ponendo il punteggio totale di *Sensation Seeking* come predittore del punteggio dell'indice β_{ott} . Il modello di regressione risulta statisticamente significativo ($F_{(1, 442)}= 11.35, p<.05$) ed il coefficiente beta standardizzato è anch'esso significativo ($t= 3.37, p<.05$) ma, come prevedibile, il modello ha un $R^2=.025$, quindi estremamente basso.

Discussione

I risultati ottenuti nella fase di messa a punto dello strumento di misurazione, hanno mostrato una corrispondenza soddisfacente tra vignette ed affermazioni, giustificando l'introduzione dei supporti visivi. Questo esito ha permesso di mettere a punto un questionario definitivo, denominato Questionario Beta, sufficientemente affidabile e comprensibile per l'utilizzo su un campione di adolescenti.

L'indice β_{ott} si è mostrato sensibile nel discriminare un diverso atteggiamento degli adolescenti rispetto alla richiesta di aiuto in base al tipo di comportamento indagato. In particolare, nonostante gli adolescenti nel complesso risultino disposti a richiedere aiuto, questo atteggiamento sembra essere ancor più marcato quando il comportamento a rischio è l'uso di droga piuttosto che l'alcool.

Questo indice ha inoltre messo in evidenza un diverso atteggiamento tra maschi e femmine rispetto alla richiesta di aiuto. Le femmine mostrano di avere un atteggiamento maggiormente rivolto alla richiesta di aiuto, mentre i maschi tendono ad adottare una strategia più conservativa e rivolta all'attesa.

In questo studio, l'indice β_{ott} non ha messo in luce un diverso atteggiamento alla richiesta di aiuto in base alla gravità delle conseguenze in cui l'adolescente può incorrere, confermando sostanzialmente che l'atteggiamento mostrato dai partecipanti è quello di richiedere aiuto sempre, indipendentemente dalla gravità delle potenziali conseguenze negative.

Infine, è emersa una relazione tra l'indice β_{ott} e la *Sensation Seeking Scale*. Coloro che ottengono dei punteggi alti a questa scala, quindi che ricercano sensazioni forti, hanno un atteggiamento più reticente a richiedere aiuto, mentre coloro che sono meno propensi alla ricerca di sensazioni risultano maggiormente disposti a chiedere aiuto. La ricerca di sensazioni, inoltre, può essere considerato un fattore predittivo della disponibilità dei giovani a richiedere aiuto.

Capitolo 4

Conclusioni

4.1 Presa di decisione in ambito medico

Per quanto concerne l'ambito di ricerca relativo alla professione medica, i risultati ottenuti nei due studi permettono di sottolineare il contributo della TDS nell'ambito della presa di decisione. Il beta ottimale risulta essere un buon metodo nello studio della valutazione dei costi e benefici nella diagnosi medica, senza condurre una analisi del comportamento medico nel contesto lavorativo, che richiederebbe altro tipo di rilevazioni.

L'influenza del fattore esperienza non è stata evidenziata dai risultati. Medici e studenti sembrano ragionare, in merito all'assunzione di rischio nella presa di decisione diagnostica, secondo il medesimo schema: è ragionevole rischiare maggiormente di fronte a patologie gravi, anche correndo il rischio di mettere il soggetto in allarme inutilmente, mentre è preferibile evitare di formulare una diagnosi a tutti i costi di fronte ad una patologia lieve, nel caso di omissioni comunque la gravità delle conseguenze è ridotta dal tipo di patologia di per sé non pericolosa.

Questi risultati sembrano in linea con alcune delle ricerche, condotte sia in fase prognostica (Dawson, 2000) che in fase diagnostica (Berwick & Thibodeau 1983; Dawson, Connors, Speroff, Kemka, Shaw & Arkes, 1993), che hanno evidenziato come il livello generale di esperienza del medico non influenzi il grado di accuratezza del giudizio.

Allo stesso tempo, i risultati ottenuti concordano con le evidenze mostrate dalla letteratura in merito alle strategie di presa di decisione in ambito non specialistico. Riteniamo, infatti, che questo tipo di atteggiamento, differenziato in base alla gravità attribuita alla situazione, possa essere considerato una modalità decisionale efficiente anche nel decisore non professionista, l'uomo comune che affronta le problematiche della vita quotidiana.

4.2 Presa di decisione in adolescenza

Relativamente ai risultati emersi dagli studi sugli adolescenti è possibile ritenere che l'indice β_{ott} possa diventare uno strumento interessante nella misurazione della disponibilità alla richiesta d'aiuto in adolescenza.

Un risultato importante, ed in linea con la letteratura, concerne il fatto che i consumatori di sostanze (specificamente *cannabis* e nicotina) hanno un approccio diverso alla richiesta di aiuto rispetto ai coetanei non consumatori: indipendentemente dalla gravità della conseguenza considerata, risultano meno propensi a richiedere aiuto. Questo risultato è in linea con la letteratura: i *participants* in comportamenti a rischio si caratterizzano per una sottostima della probabilità di incorrere in problemi, una ridotta percezione del rischio per sé stessi ed una considerazione delle possibili conseguenze come meno dannose (Benthin, Slovic & Severson, 1993; Brown, Christiansen & Goldman, 1987; Savadori & Rumiati, 1996; Slovic, 2000).

Un altro risultato ampiamente discusso in letteratura concerne le differenze di genere: in questo studio i maschi sono meno propensi delle femmine a richiedere aiuto. Questo risultato sembrerebbe mostrare da parte dei maschi una minore consapevolezza del rischio e forse una sottovalutazione della dannosità di certi comportamenti. Nonostante il problema delle differenze di genere nella propensione e valutazione del rischio in adolescenza sia controversa, questa evidenza risulta supportata da una parte della letteratura (Donohew, 1988; Greene, Krmar, Walters, Rubin & Hale, 2000; Newcomb & McGee, 1989; Tyler & Lichtenstein, 1997).

Quando si è cercato di verificare se la propensione alla richiesta di aiuto varia in base al tipo di comportamento indagato, i risultati ottenuti sono stati abbastanza omogenei: gli adolescenti sono maggiormente disposti a richiedere aiuto quando il comportamento in oggetto è l'abuso di droga, piuttosto che quando è l'abuso di alcool. In altri termini, gli adolescenti temono più la

pericolosità di conseguenze derivanti da sostanze stupefacenti che da sostanze alcoliche.

Taluni meccanismi valutativi, tipici della presa di decisione in condizioni di incertezza, possono essere chiamati in causa per tentare di spiegare questo effetto. Il consumo di alcool è un evento consueto, socialmente accettato e condiviso: nei locali pubblici, così come nella vita privata, capita spesso di vedere persone che consumano sostanze alcoliche. Il consumo di sostanze stupefacenti è un evento meno noto, socialmente condannato e di non ampia visibilità. La letteratura sulla valutazione del rischio ha mostrato che la minore familiarità con l'evento può determinare una maggiore paura dello stesso (Brun, 1994), e questo fattore potrebbe entrare in gioco nella diversa valutazione che gli adolescenti fanno di conseguenze derivanti da alcool o da droga, in particolare nel fatto che mostrano di temere di più le sostanze stupefacenti.

Un ulteriore fattore da tenere in considerazione è il controllo degli eventi: le situazioni percepite come incontrollabili sono anche giudicate come più pericolose (Bragg & Finn, 1982). Di nuovo, data la consuetudine al consumo di alcool, anche le conseguenze che ne derivano sono almeno in parte note e quindi considerate gestibili e controllabili. La droga ha degli effetti meno riconoscibili, che probabilmente sono percepiti anche come poco controllabili e quindi ingenerano più paura.

Da non trascurare infine sono gli aspetti relativi alle campagne di informazione e prevenzione che sono state, per lungo tempo, concentrate maggiormente sui problemi connessi all'utilizzo di sostanze stupefacenti nei giovani, piuttosto che all'abuso di alcool. La maggiore informazione, connessa a questo tipo di campagne, potrebbe aver sensibilizzato maggiormente i giovani rispetto al problema della droga, trascurando la reale e gravissima pericolosità delle conseguenze negative derivanti da un uso eccessivo di sostanze alcoliche. Negli ultimi anni, comunque, la tendenza è cambiata e le campagne di

informazione sui danni dell'alcool sono molto incrementate, proprio per far fronte al serio problema dei giovani e l'alcoolismo.

Per quanto concerne, infine, i risultati relativi alla sensibilità dell'indice di misurazione utilizzato rispetto alla gravità della conseguenza considerata, i risultati ottenuti nei due studi sono contrastanti e non può, al momento, essere tracciata una conclusione definitiva.

Una questione aperta, che a mio avviso influenza fortemente la misurazione mediante questo indice, è la formulazione delle domande. Per ricavare l'indice beta ottimale è necessario avere una stima soggettiva, ma sufficientemente affidabile, delle probabilità degli eventi e dei costi e benefici connessi ai possibili esiti. L'influenza della modalità di misurazione sulla possibilità di ottenere dei risultati affidabili nella percezione del rischio è nota e ampiamente studiata (Johnson & Tversky, 1984; Slovic, 2000; Tversky & Koehler, 1994; Tyszka & Goszczyńska, 1993).

Il fatto di concentrare l'attenzione su di un unico evento per volta, ad esempio, causa nel soggetto una tendenza a sovrastimare il suo presentarsi (Slovic, 2000; Tversky & Koehler, 1994) e questo è un problema che potrebbe aver influito sulla misurazione nel presente lavoro: nel primo studio gli adolescenti erano chiamati a valutare la probabilità di una serie di 9 conseguenze derivanti da comportamenti a rischio, nel secondo studio le conseguenze erano state selezionate a priori e ciascun partecipante doveva valutare la probabilità di presentarsi solo di una possibile conseguenza. La necessità di semplificare la formulazione del questionario, per facilitare la comprensibilità da parte degli adolescenti, potrebbe aver influito negativamente sull'accuratezza della stima di probabilità e, di conseguenza, sull'affidabilità dei risultati ottenuti.

Una particolare attenzione alla modalità di formulazione delle domande, permetterebbe, non solo di ottenere delle stime affidabili di probabilità degli eventi o costi-benefici degli esiti, ma anche di poter scomporre queste due componenti della valutazione e verificarne preliminarmente e separatamente

l'influenza nel processo decisionale. Un eventuale utilizzo futuro dell'indice beta ottimale anche in questo senso è certamente auspicabile e permetterebbe di studiare in modo più approfondito le caratteristiche non solo di questo tipo specifico di misura, ma dell'intero meccanismo di presa di decisione.

4.3 Conclusioni generali

Considerando nel complesso i risultati ottenuti nei due ambiti di studio prescelti (professione medica e adolescenza), l'indice beta ottimale è risultato comunque un metodo efficace per misurare le intenzioni comportamentali degli individui posti nella situazione di compiere una scelta in condizioni di incertezza. L'atteggiamento dei partecipanti rispetto all'assunzione di rischio è stato misurato efficacemente, e l'indice beta ha mostrato di variare in base a talune caratteristiche delle situazioni: questo aspetto garantisce una buona sensibilità della misurazione rispetto al fenomeno indagato. Allo stesso tempo, questa sensibilità dell'indice rispetto ad alcune variabili intervenienti può servire, indirettamente, per determinare i fattori che influenzano il processo decisionale.

Partendo dall'assunto che l'indice beta ottimale si rifà ad un approccio sperimentale al rischio e che quindi ha come obiettivo quello di fornire una misura estremamente prossima al comportamento, sarebbe interessante utilizzarlo all'interno di modelli di decisione più complessi, al fine di verificarne la collocazione e l'influenza rispetto ai fattori cognitivi, motivazionali e situazionali che, come è noto, influiscono in prima battuta sull'atteggiamento rispetto al rischio e, indirettamente, sul comportamento stesso.

Riferimenti bibliografici

- Ajzen, I. (1985). From intentions to actions: a theory of planned behavior. In J. Kuhl & J. Beckham (Eds.), *Action control: from cognition to behavior*. New York: Springer-Verlag.
- Ajzen, I. (1988). *Attitudes, personality and behavior*. Chicago, IL: Dorsey Press.
- Ajzen, I., & Fishbein, M. (1980). *Understanding attitudes and predicting social behavior*. New Jersey: Prentice-Hall.
- Arkes, H. R., Wortmann, R. L., Saville, P. D., & Harkness, A. R. (1981). Hindsight bias among physicians weighing the likelihood of diagnoses. *Journal of Applied Psychology, 66*, 252-254.
- Arnett, J. (1990a). Drunk driving, sensation seeking, and egocentrism among adolescents. *Personality and Individual Differences, 11*, 541-546.
- Arnett, J. (1990b). Contraceptive use, sensation seeking and adolescent egocentrism. *Journal of Youth and Adolescence, 19*, 171-180.
- Arnett, J. (1992). Reckless behavior in adolescence: a developmental perspective. *Developmental Review, 12*, 339-373.
- Arnett, J., & Balle-Jensen, L. (1993). Cultural bases of risk behavior: Danish adolescents. *Child Development, 64*, 1842-1855.
- Atkinson, J. M. (1957). Motivational determinants of risk taking behavior. *Psychological Review, 64*, 359-372.
- Bagozzi, R. P. (1981). Attitudes, intentions and behavior: a test of some key hypothesis. *Journal of Personality and Social Psychology, 41*, 607-627.
- Bagozzi, R. P. (1992). The self-regulation of attitudes, intentions and behavior. *Social Psychology Quarterly, 55*.
- Bagozzi, R. P., & Warshaw, P. R. (1990). Trying to consume. *Journal of Consumer Research, 17*, 127-140.
- Ball, S. A. (1995). The validity of an alternative five-factor measure of personality in cocaine abusers. *Psychological Assessment, 2*, 148-154.

- Baron, J., & Ritov, I. (1994). Reference points and omission bias. *Organizational Behavior and Human Decision Processes*, 59, 475-498.
- Benthin, A., Slovic, P., & Severson, H. (1993). A psychometric study of adolescent risk perception. *Journal of Adolescence*, 16, 153-168.
- Bentler, P. M., & Speckart, G. (1979). Models of attitude-behavior relations. *Psychological Review*, 86, 452-464.
- Bentler, P. M., & Speckart, G. (1981). Attitudes 'cause' behaviors: a structural equation analysis. *Journal of Personality and Social Psychology*, 40, 226-238.
- Berwick, D. M., & Thibodeau, L. A. (1983). Receiving operating characteristics analysis of diagnostic skill. *Medical Care*, 21, 876-885.
- Blackwell, H. R. (1952). The influence of data collection procedures upon psychophysical measurement of two sensory functions. *Journal of Experimental Psychology*, 44(5), 306-315.
- Blackwell, H. R. (1963). Neural theories of simple visual discrimination. *Journal of the Optical Society of America*, 53, 129-160.
- Bragg, B. W. E., & Finn, P. (1982). *Young Driver Risk-Taking Research: Technical Report of Experimental Study*. Washington: Department of Transportation.
- Brun, W. (1994). Risk perception: main issues, approaches and findings. In G. Wright & P. Ayton (Eds.), *Subjective Probability* (pp. 295-320). London: Wiley.
- Caplan, R. A., Posner, K. L., & Cheney, F. W. (1991). Effect of outcome on physician judgements of appropriateness of care. *Journal of the American Medical Association*, 265, 1957-1960.
- Chapman, G. B., & Sonnenberg, F. A. (2000). *Decision making in health care: theory, psychology and applications*. New York: Cambridge University Press.

- Chapmann, G. B., & Elstein, A. S. (2000). Cognitive processes and biases in medical decision making. In G. B. Chapman & F. A. Sonnenberg (Eds.), *Decision making in health care: theory, psychology and applications* (pp. 183-210). New York: Cambridge University Press.
- Chassin, L. (1985). Changes in peer and parent influence during adolescence: Longitudinal versus cross-sectional perspectives on smoking initiation. *Developmental Psychology, 22*, 327-334.
- Clasen, D., & Brown, B. (1985). The multidimensionality of peer pressure in adolescence. *Journal of Youth and Adolescence, 14*, 451-468.
- Collins, J. K., & Harper, J. F. (1985). Sexual behavior and peer pressure in adolescent girls. *Australian Journal of Sex, Marriage, and Family, 6*, 137-142.
- Connors, A. F., Dawson, N. V., Speroff, T., Arkes, H., Knaus, W. A., Harrel, F. E., Lynn, J., Terro, J., Goldman, L., Califf, R., Fulkerson, W., Oye, R., Bellamy, P., & Desbiens, N. (1992). Physicians' confidence in their estimates of the probability of survival: relationships to accuracy. *Medical Decision Making, 12*, 336.
- Damody, J. (1991). The adolescent personality, formal reasoning, and values. *Adolescence, 26*, 731-742.
- Dawson, N. V. (2000). Physician judgement of uncertainty. In G. B. Chapman & F. A. Sonnenberg (Eds.), *Decision making in health care: theory, psychology and applications* (pp. 210-252). New York: Cambridge University Press.
- Dawson, N. V., Connors, A. F., Speroff, T., Kemka, A., Shaw, P., & Arkes, H. R. (1993). Hemodynamic assessment in managing the critically ill: is physician confidence warranted? *Medical Decision Making, 13*, 258-266.
- Dellarosa, D., & Bourne, L. E. (1984). Decisions and memory: differential retrievability of consistent and contradictory evidence. *Journal of Verbal Learning and Verbal Behavior, 23*, 669-682.

- Dishion, T. J., & Loeber, R. (1985). Adolescent marijuana and alcohol use: the role of parents and peers revisited. *American Journal of Drug and Alcohol Use, 11*, 11-25.
- Donohew, L. (1988). Effects of drug abuse message styles: final report, grant no. DA03462. Rockville, MD: National Institute on Drug Abuse.
- Drottz-Sjoberg, B. M. (1991). *Perception of risk. Studies of risk attitudes, perceptions and definitions*. Stockholm: Stockholm School of Economics, Center for Risk Research.
- Edwards, W. (1992). *Utility theories: measurements and applications*. Boston: Kluwer Academic Publishers.
- Einhorn, H. J., & Hogarth, R. M. (1978). Confidence in judgment: persistence of the illusion of validity. *Psychological Review, 85*, 395-416.
- Elkind, D. (1967). Egocentrism in adolescence. *Child Development, 38*, 1025-1034.
- Fechner, G. T. (1860). *Elemente der Psychophysik*. Leipzig: Breitkopf & Hartel (English translation: Adler, H. E. (1966). *Elements of Psychophysics*. New York: Holt, Rinehart & Winston).
- Finley, J. P. (1884). Tornado Predictions. *American Meteorological Journal, 1*, 5-88.
- Fischhoff, B. (1975). Hindsight \neq foresight: the effect of outcome knowledge on judgement under uncertainty. *Journal of Experimental Psychology: Human Perception and Performance, 1*, 288-299.
- Fishbein, M., & Ajzen, I. (1975). *Belief, attitude, intention and behavior*. Reading, MA: Addison-Wesley.
- Fishbein, M., Middlestadt, S. E., & Hitchcock, P. J. (1994). Using information to change sexually transmitted disease-related behaviors. In R. J. DiClemente & J. L. Peterson (Eds.), *Preventing AIDS: theories and methods of behavioral interventions* (pp. 61-78). New York: Plenum Press.

- Fischer, G. W., Morgan, M. G., Fischhoff, B., Nair, I. & Lave, L. B. (1991). What risks are people concerned about? *Risk Analysis*, *11*, 303-314.
- Fisk, A. D., & Schneider W. (1984). Memory as a function of attention, level of processing, and automatization. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *10*, 181-197.
- Fredricks, A. J., & Dosset, D. L. (1983). Attitude behavior relations: a comparison of the Fishbein-Ajzen and Bentler-Speckart models. *Journal of Personality and Social Psychology*, *45*, 501-512.
- Galambos, N. L., & Silbereisen, R. K. (1987). Substance use in West Germany youth: a longitudinal study of adolescents' use of alcohol and tobacco. *Journal of Adolescent Research*, *2*, 161-174.
- Gillund, G., & Shiffrin, R. M. (1984). A retrieval model for both recognition and recall. *Psychological Review*, *91*, 1-67.
- Goodman, L. A. (1970). The multivariate analysis of qualitative data: Interactions among multiple classifications. *Journal of the American Statistical Association*, *45*, 226-256.
- Green, D. M. (1964). General prediction relating yes-no and forced choice results. *Journal of the Acoustical Society of America*, *36*, 1042.
- Green, D. M., & Swets, J. A. (1966). *Signal detection theory and psychophysics*. New York: Wiley.
- Greene, K., Krcmar, M., Walters, L. H., Rubin, D. L., & Hale, J. (2000). Targeting adolescent risk-taking behaviors: the contributions of egocentrism and sensation-seeking. *Journal of Adolescence*, *23*, 439-461.
- Greening, L., & Stoppelbein, L. (2000). Young drivers' health attitudes and intentions to drink and drive. *Journal of Adolescent Health*, *27*, 94-101.
- Grier, J. B. (1971). Nonparametric indexes for sensitivity and bias: Computing formulas. *Psychological Bulletin*, *74*, 424-429.

- Hampson, S. E., Severson, H. H., Burns, W. B., Slovic, P., & Fisher, K. J. (2001). Risk perception, personality factors and alcohol use among adolescents. *Personality and Individual Differences, 30*, 167-181.
- Hansen, W. B., Raynor, A. E., & Wolkenstein, B. H. (1991). Perceived personal immunity to the consequences of drinking alcohol: the relationship between behavior and perception. *Journal of Behavioral Medicine, 14*, 205-224.
- Hays, W. L. (1973). *Statistics for the social sciences* (2nd ed). New York: Holt, Rinehart & Winston.
- Hodos, W. (1970). Nonparametric index of response bias for use in detection and recognition experiments. *Psychological Bulletin, 74*, 351-354.
- Ingham, J. G. (1970). Individual differences in signal detection. *Acta Psychologica, 34*, 39-50.
- Inhelder, B., & Piaget, J. (1958). *The growth of logical thinking from childhood to adolescence*. New York: Basic Books.
- Jessor, R. (1984). Adolescent development and behavioral health. In J. D. Matarazzo, S. M. Weiss, J. A. Herd & N. E. Miller (Eds.), *Behavioral health: a handbook of health enhancement and disease prevention* (pp. 69-90). New York: Wiley.
- Jessor, R. (1990). Road safety and health behavior: some lessons for research and intervention. *Health Education Research, 5*, 1-3.
- Jessor, R., & Jessor, S. L. (1975). Adolescent development and the onset of drinking. *Journal of Studies if Alcohol, 36*, 27-51.
- Jessor, R., & Jessor, S. L. (1977). *Problem-behavior and psychosocial development: a longitudinal study of youth*. New York: Academic.
- Johnson, B. B. (1993). Advancing understanding of knowledge's role in lay risk perception. *RISK: Issues in Health and Safety, 4*, 189-212.
- Johnson, E. J., & Tversky, A. (1984). Representations of perceptions of risks. *Journal of Experimental Psychology: General, 113*, 55-70.

- Kahneman, D., & Tversky, A. (1979). Prospect theory: an analysis of decision under risk. *Econometrica*, *47*, 263-291.
- Kahneman, D., & Tversky, A. (1984). Choices, values and frames. *American Psychologist*, *39*, 341-350.
- Kalmuss, D. (1986). Contraceptive use: a comparison of ever- and never-pregnant adolescents. *Journal of Adolescent Health Care*, *7*, 332-337.
- Landis, J. R., & Koch, G. C. (1977). The measurement of observer agreement for categorical data. *Biometrics*, *33*, 159-174.
- Leone, L., Perugini, M., & Ercolani, A. P. (1999). A comparison of three models of attitude-behavior relationships in the studying behavior domain. *European Journal of Social Psychology*, *29*, 161-189.
- Lopes, L. L. (1987). Between hope and fear: The psychology of risk. *Advances in Experimental Social Psychology*, *20*, 255-295.
- Luce, R. D. (1959). *Individual choice behaviour*. New York: Wiley.
- Luce, R. D. (1963). Detection and recognition. In R. D. Luce, R. R. Bush, & E. Galanter (Eds.), *Handbook of mathematical psychology* (pp. 103-189). New York: Wiley.
- Macmillan, N. A., & Creelman, C. D. (1990). Response bias: characteristics of Detection Theory, Threshold Theory, and “Nonparametric” Indexes. *Psychological Bulletin*, *107*, 401-413.
- Macmillan, N. A., & Creelman, C. D. (1996). Triangles in ROC space: History and theory of “nonparametric” measures of sensitivity and response bias. *Psychonomic Bulletin & Review*, *3*, 164-170.
- McNicol, D. (1972). *A primer of signal detection theory*. London: Allen & Unwin.
- Moore, S., & Gullone, E. (1996). Predicting adolescent risk behavior using a personalized cost-benefit analysis. *Journal of Youth and Adolescence*, *25*, 343-359.

- Newcomb, M. D., & McGee, L. (1989). Adolescent alcohol use and other delinquent behaviors: a one year longitudinal analysis controlling for sensation seeking. *Criminal Justice and Behavior*, *16*, 345-369.
- Neyman, J., & Pearson, E. S. (1933). On the Problem of the Most Efficient Tests of Statistical Hypotheses. *Philosophical Transactions of the Royal Society of London, Ser. A*, *231*, 289-337.
- Pastore, R. E., Crawley, E. J., Berens, M. S., & Skelly, M. A. (2003). "Nonparametric" A' and other modern misconceptions about signal detection theory. *Psychonomic Bulletin & Review*, *10*(3), 556-569.
- Peterson, W. W., Birsdall, T. G., & Fox, W. C. (1954). The theory of signal detectability. *Transactions of the IRE Professional Group on Information Theory, PGIT-4*, 171-212.
- Pollack, I., & Norman, D. A. (1964). A nonparametric analysis of recognition experiments. *Psychonomic Science*, *1*, 125-126.
- Redelmeier, D. A., Koehler, D. J., Liberman, V., & Tversky, A. (1995). Probability judgement in medicine: discounting unspecified possibilities. *Medical Decision Making*, *15*(3), 227-230.
- Savadori, L., & Rumiati, R. (1996). Percezione del rischio negli adolescenti italiani. *Giornale Italiano di Psicologia*, *1*, 85-105.
- Savage, L. J. (1954). *The foundations of statistics*. New York: Wiley.
- Schwarz, R. M., Burkhart, B. R., & Green, S. B. (1978). Turning on or turning off? Stimulus seeking or tension reduction in drinking behavior. *Journal of Consulting and Clinical Psychology*, *46*(5), 1144-1145.
- Shafir, E., Simonson, I., & Tversky, A. (1997). Reason-based choice. In W. M. Goldstein & R. M. Hogarth (Eds.), *Research on Judgement and Decision Making: Currents, Connections, and Controversies* (pp. 69-94). New York: Cambridge University Press.
- Sheer, V. C., & Cline, R. J. (1994). The development and validation of a model explaining sexual behavior among college students: Implications for

- AIDS communication campaigns. *Human Communication Research*, 21, 280-304.
- Simon, H. A. (1956). Rational choice and the structure of the environment. *Psychological Review*, 63, 129-138.
- Simon, H. A. (1959). Theories of decision making in economics and behavioral science. *American Economic Review*, 49, 253-283.
- Sjoberg, L., & Torell, G. (1993). The development of risk acceptance and moral valuation. *Scandinavian Journal of Psychology*, 34, 223-236.
- Sjoberg, L., & Winroth, E. (1986). Risk, moral value of actions, and mood. *Scandinavian Journal of Psychology*, 27, 191-208.
- Slovic, P. (1964). Assessment of risk taking behavior. *Psychological Bulletin*, 61(3), 220-233.
- Slovic, P. (1987). Perception of risk. *Science*, 236, 280-285.
- Slovic, P. (1995). The construction of preference. *American Psychologist*, 50, 364-371.
- Slovic, P. (2000). *The perception of risk*. London: Earthscan Publications Ltd.
- Slovic, P., & Lichtenstein, S. (1983). Preference reversals: a broader perspective. *American Economic Review*, 73, 596-605.
- Slovic, P., Fischhoff, B., & Lichtenstein, S. (1980). Facts and fears: understanding perceived risk. In R. Schwing & W. A. Alberts (Eds.), *Societal risk assessment: how safe is safe enough?* (pp. 181-216). New York: Plenum Press.
- Slovic, P., Lichtenstein, S., & Fischhoff, B. (1984). Modeling the social impact of fatal accidents. *Management Science*, 30, 464-474.
- Snodgrass, J. G., & Corwin, J. (1988). Pragmatics of measuring recognition memory: Applications to dementia and amnesia. *Journal of Experimental Psychology: General*, 117, 34-50.

- Stanislaw, H., & Todorov, N. (1999). Calculation of signal detection theory measures. *Behavior Research Methods, Instruments, & Computers*, *31*(1), 137-149.
- Statistical Package for the Social Sciences. (2000). *SPSS for Windows 11.5*. Chicago: SPSS Inc.
- Swets, J. A. (1964). Signal detection and recognition by human observers. New York: John Wiley & Sons.
- Swets, J. A. (1986). Indices of discrimination or diagnostic accuracy: their ROCs and implied models. *Psychological Bulletin*, *99*, 100-117.
- Swets, J. A. (1996). Signal detection theory and ROC analysis in psychology and diagnostics. Collected Papers. Mahwah, New Jersey: Lawrence Erlbaum Associates.
- Swets, J. A., & Pickett, R. M. (1982). Evaluation of diagnostic systems: Methods from signal detection theory. New York: Academic Press.
- Tanner, W. P., & Swets, J. A. (1954a). The human use of information, I: Signal detection for the case of the signal known exactly. *Transactions of the IRE Professional Group on Information Theory, PGIT-4*, 213-221.
- Tanner, W. P., & Swets, J. A. (1954b). A decision-making theory of visual detection. *Psychological Review*, *61*, 401-409.
- Teigen, K. H., Brun, W., & Slovic, P. (1988). Societal risk as seen by a Norwegian public. *Journal of Behavioral Decision Making*, *1*, 111-130.
- Thurstone, L. L. (1927). A law of comparative judgement. *Psychological Review*, *34*, 273-286.
- Tversky, A., & Kahneman, D. (1973). Availability: a heuristic for judging frequency and probability. *Cognitive Psychology*, *5*, 207-232.
- Tversky, A., & Koehler, D. J. (1994). Support theory: a nonextensional representation of subjective probability. *Psychological Review*, *101*(4), 547-567.

- Tyler, J., & Lichtenstein, C. (1997). Risk, Protective, AOD Knowledge, Attitude, and AOD Behavior. Factors Associated with Characteristics of High-Risk Youth. *Evaluation and Program Planning*, 20(1), 27-45.
- Tyszka, T., & Goszczynsli, M. (1993). What verbal reports say about risk perception. *Acta Psychologica*, 83, 53-64.
- Von Neumann J., & Morgenstern O. (1947). *Theory of Games and Economic Behavior*. Princeton: Princeton University Press.
- Wald, A. (1950). *Statistical Decision Functions*, New York: Wiley.
- Woodcock, F. (1976). The evaluation of yes-no forecasts for scientific and administrative purposes. *Monthly Weather Review*, 104, 1209-1214.
- Woodworth, R. S. (1938). *Experimental psychology*. New York: Holt, Rinehart & Winston.
- Yates, F. J., & Stone, E. R. (1992). The risk construct. In F. J. Yates (Ed.), *Risk Taking Behavior* (pp. 1-25). London: Wiley.
- Yule, G. U. (1912). On the methods of measuring association between two attributes. *Journal of the Royal Statistical Society*, 75, 579-642.
- Zuckerman, M. (1978). Sensation Seeking. In H. London & J. Exner (Eds.), *Dimensions of Personality* (pp. 487-559). New York: Wiley.
- Zuckerman, M. (1979). *Sensation seeking: beyond the optimal level of arousal*. Hillsdale, NJ: Erlbaum.
- Zuckerman, M., & Neeb, M. (1980). Demographic influences in sensation seeking and expressions of sensation in religion, smoking, and driving habits. *Personality and Individual Differences*, 1, 197-206.