LA MATEMATICA DELLA SCALOGNA.

Prerequisiti: le quattro operazioni

L'affermazione originale della famosa "Legge di Murphy" è:"Se qualcosa può andar male, lo farà". Molti, guardandosi intorno, aggiungerebbero "…a me in particolare". E' il modo sbagliato di guardare il mondo.   In molti casi, se le cose vanno in un modo irritante è semplicemente perché la probabilità che vadano come ci piace è inferiore a quella che vadano tutto al contrario. Non si può dire di essere sfortunati se le probabilità non ci sono favorevoli: in altre parole, non è sfortunato chi non vince all'Enalotto. Semmai è molto fortunato chi vince.

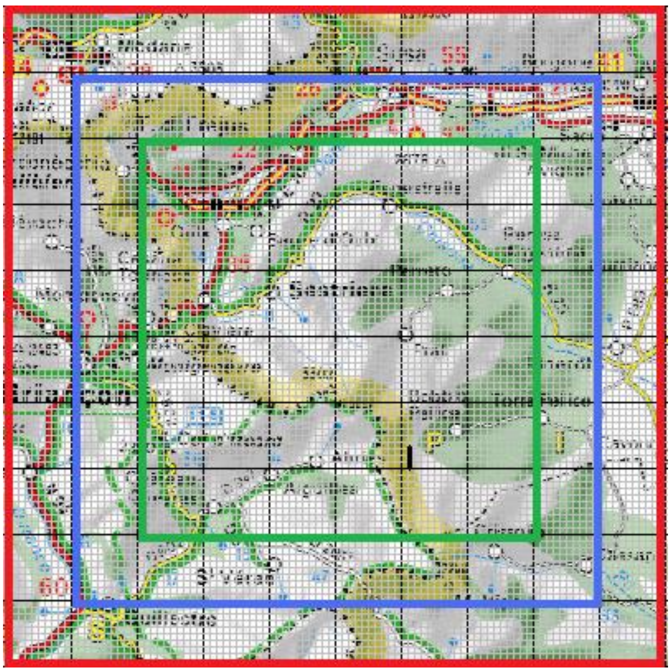
Vorrei fare alcuni esempi:

1) Le code al supermercato o in autostrada: "quella che scelgo io è sempre la più lenta".

Se in un supermercato voi scegliete alla cieca un'uscita su 10 possibili, ciascuna con la sua brava cassiera ed i clienti che aspettano con carrelli variamente carichi, e con code tutte più o meno lunghe eguali, avete solo una probabilità su dieci di aver scelto la coda che si dimostrerà più veloce. Ci dovrebbe consolare pensare che nove su dieci clienti non sono nella coda più veloce, esattamente come noi.   Allo stesso modo, in un'autostrada a tre corsie per direzione abbiamo – in generale - solo una probabilità su tre di essere nella corsia più veloce rispetto alle due che ci affiancano. Di nuovo, due automobilisti su tre sono infelici pur non essendo sfortunati.   Uguali, poi, sono le probabilità di aver scelto la coda o la corsia più lenta.

2) Le carte topografiche. "Quando sono in viaggio, la zona di una carta stradale che mi interessa è sempre sul margine".

Anche questo è dovuto alla probabilità che ci sono contrarie nonostante a vista sembrino favorevoli o almeno neutre.



Se si sceglie un punto a caso nella mappa, la probabilità che esso si trovi in una data zona della mappa è data dal rapporto fra l'area della zona e l'area totale. Valutare l'area a occhio è ingannevole. L'area compresa fra il quadrato blu e quello rosso è 36%, più di un terzo del totale. L'area compresa fra il quadrato verde e quello rosso è il 64%, quasi due terzi. Quindi abbiamo un poco più di una possibilità su tre di essere all'interno del quadrato verde, e un po' meno di due su tre di essere all'interno del quadrato blu. Non abbastanza da gridare alla scalogna se ci troviamo sul bordo.

3) Se una fetta di pane imburrato cade, tende a cadere sulla faccia imburrata (ragionamento che apparve sullo Scientific American una ventina d'anni fa).

Il fatto fu già notato in epoca vittoriana, quando evidentemente tutti prendevano il pane imburrato col tè e l'incidente poteva essere considerato una catastrofe. Qui le ragioni sono assai più sottili. La fetta modello (che supporremo piatta, di pane tostato) cade normalmente più o meno da un'altezza che è quella dei nostri tavoli. Questa a sua volta è legata alla nostra altezza, la quale, a sua volta, non è casuale, essendo legata all'equilibrio tra la resistenza delle ossa e la forza di gravità. La resistenza delle ossa è dettata dalle forze elettrostatiche che tengono lontani i nostri atomi e molecole e, in ultima analisi, dalla costante elettromagnetica che determina la forza delle interazioni elettromagnetiche. La faccia imburrata è normalmente la faccia superiore. Quando la fetta cade dal tavolo, per poco che non cada bella piatta, esperimenti fatti anche con altri oggetti piatti di simile geometria hanno dimostrato che compie circa mezzo giro e così la faccia superiore si trova di sotto. Il mezzo giro è legato alla relazione tra l'altezza del tavolo e la forza di gravitazione alla superficie della terra, cioè, in ultima analisi, alla costante gravitazionale. Il rapporto tra le due costanti, elettromagnetica e gravitazionale, è un dato nella struttura del nostro universo. Insomma, la caduta sulla faccia imburrata è un fatto universale, progettato alle origini dell'universo, e considerarla una nostra particolare scalogna è come pensare di essere particolarmente scalognati perché d'inverno è più freddo che d'estate. A parte il fatto che un gatto cade sempre sulle quattro zampe, e c'è chi si chiede che cosa succederebbe ad imburrare la schiena di un gatto e poi farlo cadere. Alcuni parlano di moto perpetuo…

4) Ho rinunciato a fare una gita domenica perché era predetto brutto tempo, e invece è stata una giornata bellissima.

Questo succede più sovente di quanto non si creda, soprattutto perché in generale non è ben chiaro il discorso dell'attendibilità delle previsioni. Supponiamo che nella zona dove vogliamo fare la gita faccia brutto sessanta giorni all'anno (lasciamo perdere le sfumature di bello e di brutto tempo, come il nuvoloso, il variabile etc.) . Supponiamo inoltre che la previsione del tempo sia attendibile al 90%. Ciò vuol dire che il servizio meteorologico in circa 54 giorni all'anno (90% di 60) prevede brutto tempo quando poi è effettivamente brutto tempo ed in circa 30 giorni all'anno (10% dei restanti 300 giorni dell'anno) prevede brutto tempo quando poi c'è sole. In conclusione prevede brutto tempo 84 giorni all'anno, ma solo in 54 di essi effettivamente fa brutto. Cioè, quando il servizio meteorologico ci dice che farà brutto, la probabilità di brutto tempo non è del 90 per cento come avevamo creduto, ma solo 54/84, cioè circa 2 su 3. Di nuovo, non si può gridare alla scalogna se il tempo è poi stato bello. Da notare che quanto più pochi sono i giorni di brutto tempo, tanto più alta è la probabilità che faccia bello quando viene previsto brutto tempo (provate con altre percentuali). Non è poi così strano. In questo come in molti altri campi, un "tasso di successo" del 90%, per quanto apparentemente affidabile, non significa quello che sembra.

Riassumendo:

Dunque per le fette imburrate, la caduta sul lato sbagliato dipende da come è disegnato l'universo. Per i bordi delle mappe e per il bel tempo le probabilità "sulla carta" non sono favorevoli come sembrano. Per le code ci sono diverse possibili contromisure: la più ovvia è quella di cercare di valutare la lunghezza della coda non in numero di clienti ma in numero totale di oggetti che la cassiera dovrà registrare. Qui la fortuna non c'entra. E lo stesso vale per le corsie nelle autostrade. L'esperienza può insegnare quali corsie vanno evitate a seconda delle regole di circolazione, delle uscite e degli ingressi in autostrada, dell'ora etc. Dunque sì, ci sono i fortunati che si trovano per caso nella coda più veloce, ma nelle code più veloci ci sono anche quelli che pensano, il che non ha nulla a che fare con la fortuna. Tutti gli altri non sono particolarmente sfortunati, semplicemente non sono fortunati. Ma tra le due cose c'è un mare di differenza.

NOTA 1 (sul "brutto tempo").

La spiegazione del problema del brutto tempo aiuta a risolvere un altro problema, che ha meno a che vedere con la scalogna. Supponiamo di avere due scatole di dolci. In una (scatola A) ci sono10 cioccolatini e 30 gelatine di frutta, nell'altra (scatola B) ci sono 20 cioccolatini e 20 gelatine di frutta. Senza guardare la scatola prendiamo un dolce. E' una gelatina di frutta. Qual è la probabilità che abbiamo pescato dalla scatola A? A priori la probabilità di pescare da una delle due scatole è la stessa, ma adesso abbiamo qualche informazione in più. Il problema diventa ovvio se viene formulato in modo lievemente diverso: qual è la probabilità che la gelatina di frutta venga dalla scatola A? La risposta è che in tutto ci sono 50 gelatine di frutta, di cui 30 nella scatola A , cioè 30/50 del totale, e 20 nella scatola B. Dunque la probabilità che la gelatina venga dalla scatola A è 3/5, quella che venga dalla scatola B è 2/5. E questo risponde anche alla domanda originale, di poco differente, "qual è la probabilità che abbiamo pescato dalla scatola A?". Quindi la probabilità a priori che era ½ di pescare in una delle due scatole, è diventata - a posteriori - 3/5 a vantaggio della A grazie al fatto nuovo, che abbiamo pescato una gelatina di frutta. La cosa diventerebbe naturalmente ovvia se ci fosse in tutto una sola gelatina di frutta, e fosse nella scatola A. In questo caso la probabilità che la gelatina di frutta provenga dalla scatola A è 1, certezza.
Questa tecnica, più o meno adattata a casi anche assai più complessi, ha larga applicazione in medicina, ad esempio nella diagnosi di malattie. Se riuscite positivi a un test (sicuro al 90%) se avete l'influenza, in base al ragionamento fatto, non significa che avete l'influenza col 90% di probabilità. Dipende anche da quanto è diffusa l'influenza. In ogni caso avrete l'influenza **al massimo** col 90% di probabilità ed in molti casi sarà assai meno probabile. Quindi, se sperate di avere l'influenza per evitare il compito in classe di matematica, non contateci troppo, tanto più che se tutta la classe ha l'influenza il compito in classe viene normalmente rimandato.

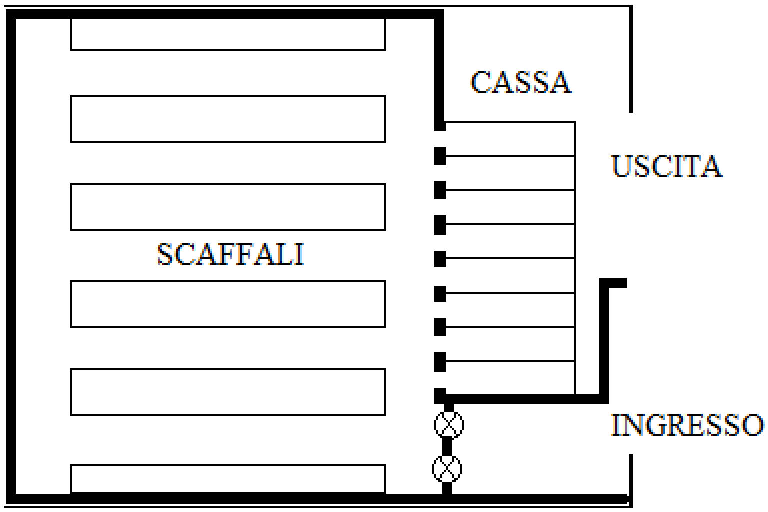
> Vediamo il conto per sicurezza. Supponiamo che 10% degli scolari abbia l'influenza e il test sia preciso al 90%. Vuol dire che il test sbaglierà una diagnosi su 10 tra quelli che hanno l'influenza (cioè dirà che sono 9 ad averla), e 9 sui rimanenti novanta scolari, cioè dirà che hanno l'influenza mentre non l'hanno. In altre parole prevederà 18 ammalati, mentre lo saranno solo la metà. Se siete riuscito positivo, la probabilità di avere effettivamente l'influenza è del 50%.

NOTA 2 (sull'attesa in coda).

Nessuno si stupirà ad apprendere che esiste una disciplina matematica che si chiama "Teoria delle code" , la quale ha un suo interesse e può raggiungere gradi notevoli di difficoltà.
Io mi limiterò a presentare un teorema non banale ma molto generale [[John Little, 1961]], che ci può illuminare su un certo numero di questioni, e che io presenterò nella forma e nel modo più semplice, forse addirittura corretto.

Consideriamo uno schema di SUPERMERCATO.
Non tutti i supermercati sono fatti come quello del mio modello, ma questo va bene per i nostri scopi. Inoltre, non sono solo i supermercati ad avere questo schema. Grandi negozi, banche, biglietterie di stazioni ferroviarie, autostrade, aereoporti , natalità e mortalità in una città o Paese, ed altro in ultima analisi funzionano così.
I clienti entrano dall'INGRESSO, restano nell'area SCAFFALI per un certo tempo, poi passano dalle varie code di CASSA ed escono attraverso l'USCITA. Le maiuscole indicano che la parola può applicarsi a qualsiasi altro sistema analogo, facendo i dovuti cambiamenti.

Supponiamo di metterci vicino all'ingresso, che per comodità mettiamo vicino all'uscita. Contiamo approssimativamente quanti clienti entrano ogni minuto, per esempio 6. A noi interessa un numero medio, quindi contiamo quanti clienti entrano in dieci minuti e poi dividiamo per dieci. Noto che se ci sono madre e figlio poppante insieme, il poppante non conta. Per due adulti dipende, se ci aspettiamo che si separino quando vagano per gli scaffali, nel qual caso li contiamo come due clienti, o se probabilmente staranno sempre insieme, nel qual caso contano per un solo cliente. E ci sono ancora diverse possibilità intermedie. Ma non ci complichiamo troppo la vita.



Ora andiamo all'uscita senza aspettare troppo tempo e seguendo le stesse regole contiamo quanta gente esce ogni minuto.

Adesso ci sono tre possibilità.
1) Entra più gente di quanta esca. Non possiamo parlare di una situazione "stazionaria", in cui per esempio il numero di clienti dentro al supermercato (che chiameremo N) è costante nel tempo anche se i clienti non sono sempre le stesse persone.
   Siamo per esempio in fase di apertura, ed N aumenta.
2) Entra tanta gente quanta ne esce, nel qual caso siamo in un caso "stazionario", e possiamo aspettarci che il numero di clienti dentro al supermercato sia più o meno costante.
3) Esce più gente di quanta ne entri. Di nuovo siamo in situazione "non-stazionaria", ed il numero dei clienti all'interno decresce.

Il teorema di Little si applica solo al caso (2). Non è un disastro, perché negli altri due casi N cresce o decresce e quindi, a meno che non si vogliano fare studi molto dettagliati, interessa meno sapere quanto vale N ad un dato istante, visto che un minuto dopo sarà diverso.

Dunque la situazione è come segue:

(Persone che entrano in SUPERMERCATO ogni minuto) = (persone che entrano nell'area SCAFFALI ogni minuto) = (persone che escono da SCAFFALI ed entrano in CASSA ogni minuto) = (persone che escono da CASSA e da SUPERMERCATO ogni minuto).

Se queste eguaglianze non valessero, avremmo un problema. Per esempio, se entrassero in SUPERMERCATO più persone di quante ne escono da USCITA avremmo una montagna crescente di clienti (o più di una montagna) da qualche parte.
Ora il teorema di Little è difficile da dimostrare, ma facile da enunciare e relativamente facile da illustrare. Scegliamo la seconda via.
Questo è l'enunciato: in un caso "stazionario" come quello descritto, il numero N di clienti all'interno di SUPERMERCATO è dato da una semplice espressione, in cui D è l'intervallo medio di ingresso tra i clienti (nel nostro caso ne entrano 6 al minuto e quindi l'intervallo è 1/6 =0.167 minuti, 10 secondi) e T è il tempo di permanenza dei clienti in SUPERMERCATO. Tanto T quanto D devono essere misurati nelle stesse unità. L'espressione è:

$$N = T/D$$

Da cui derivano:

$$D = \frac{T}{N} \ e \ \ anche \ \ T \ = N \ \cdot D$$

Se ci viene più facile, in luogo della variabile D possiamo usare la variabile r, il numero di clienti che entra per minuto. Come abbiamo visto, questo non è altro che

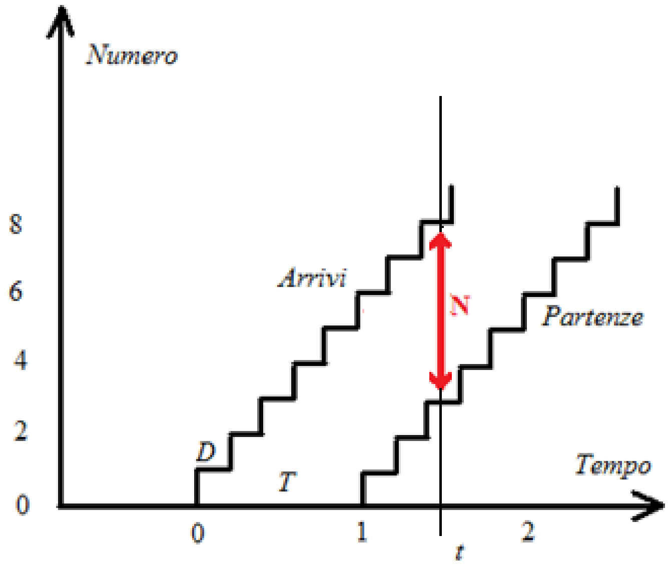
$$r = \frac{1}{D} clienti/minuto$$

Da cui:

$$N = r \ (clienti\ /minuto) \times T \ (minuti).$$

Illustrazione.

Supponiamo un caso elementare, in cui arriva in un negozio un cliente esattamente ogni due minuti e ogni cliente sta esattamente 10 minuti nel negozio. Nei primi dieci minuti il negozio si popola: cinque persone entrano prima che il primo esca. Dopodiché, per ogni cliente che esce ce n'è uno che entra e il numero di presenti resta costante. Questo numero è quello che si è costruito nel primo intervallo T di dieci minuti, perché in quel tempo sono entrati cinque clienti e nessuno è uscito. Naturalmente, 5 è dato da r · T = T/D , dove T = 10 e r = 0.5, ovvero D = 2. Forse la figura sottostante può aiutare:



La popolazione N al tempo t è data dalla differenza tra la linea degli arrivi (cumulativi) e quella delle partenze (cumulative) in quell'istante t. In figura, a partire da t = 1, è costante e vale 5.

Potete divertirvi a fare altri diagrammi a scaletta. Per esempio, ad un certo punto potete aumentare r, per esempio raddoppiandolo. Vedrete che N raddoppierà, impiegando un tempo T per farlo. Oppure potreste raddoppiare T passando a 2T. Anche qui, vedrete che N raddoppierà, impiegando il tempo 2T, che è il nuovo tempo di permanenza nel SUPERMERCATO.

Applicazioni.

Sul tempo medio di permanenza T certamente i SUPERMERCATI hanno statistiche dettagliate . Mettiamo che, inclusa l'attesa nelle code di CASSA, in un dato periodo di un dato giorno della settimana, T sia 20 minuti. Sulla base di r = 6 (cioè che entrano 6 clienti al minuto), questo vuol dire che nel SUPERMERCATO ci sono 120 clienti.

Se invece non conosciamo T, possiamo trovarlo contando quanti clienti ci sono nel supermercato. Supponiamo di trovarne 240, da cui risulta che il tempo medio di permanenza è 240/6 = 40 minuti. Magari è sabato, la gente compra per il week-end, e ci sono lunghe code alle uscite.

Applichiamo il teorema anche alle code nell'area di CASSA. Qui, considerando lo stesso caso, entrano ed escono ancora 6 clienti al minuto: se ne entrano di più di quanti ne escano, le code aumentano indefinitamente di lunghezza ed il manager del SUPERMERCATO farà bene ad aprire altri sportelli di cassa prima di essere linciato dai clienti.
Nell'area di CASSA entrano dunque 6 clienti al minuto, e ne devono uscire altrettanti.
Il manager, che conosce il tempo medio che occorre ad un cassiere per sbrigare un cliente, ha probabilmente messo al lavoro un numero di cassieri tale che tanti clienti entrino nell'area di CASSA, cioè nel SUPERMERCATO, quanti ne escono dall'area di CASSA. Questo lo ha fatto quanto le code hanno incominciato ad allungarsi fuori controllo, ed hanno raggiunto un certo numero medio di membri, che incominciava a destare nervosismo, per esempio 5 (numero, notate bene, in questo caso arbitrario e dettato dall'esperienza e dalla situazione).

Una volta che il provvedimento di aggiungere cassieri è stato preso, se le condizioni non cambiano, il numero medio di persone in coda è N=5, e non muterà più. Supponiamo che il tempo per sbrigare un cliente sia di 2 minuti. In altre parole viene sbrigato mezzo cliente al minuto. Vale la relazione:

(Numero di cassieri) × (0.5 clienti sbrigati al minuto) = 6 clienti al minuto.

Quindi, per far uscire 6 clienti al minuto evidentemente dobbiamo avere 12 cassieri. Quando un cliente entra in coda, ha davanti a sé cinque persone, e il suo tempo di attesa sarà in media 2 minuti per cliente, 10 minuti in tutto.
Nell'area di CASSA ci sono ora 5 persone per ciascuna delle 12 code, per un totale di 60 persone, il che rispetta il teorema di Little, N = r T = 6 clienti/minuto × 10 minuti.
Dunque in questo caso, del tempo totale di permanenza nel SUPERMERCATO, solo 10 minuti passano aspettando in coda alla cassa.

Contando il numero di casse aperte ed il numero di clienti che entrano al minuto nel SUPERMERCATO possiamo sapere quanto tempo occorre per sbrigare un cliente medio da parte di un cassiere medio. Moltiplicando questo tempo per il numero medio di persone in coda per ogni coda, possiamo anche stimare quanto tempo dovremo attendere in coda, a meno di essere particolarmente fortunati/sfortunati – o a meno che il manager decida di aprire o chiudere altre casse.

Si vedono qui anche talune strategie che un SUPERMERCATO può adottare. Se, avendo già il locale normalmente affollato al limite della capienza, il manager desidera avere ancor più clienti in un giorno (quindi vuol accrescere r) deve, o aumentare l'area degli SCAFFALI per poter ricevere più persone, o diminuire T, il tempo di permanenza. Quindi, se mentre state bighellonando tra gli SCAFFALI vi si avvicina un impiegato che vi chiede cortesemente: "Posso aiutarla?", sapete in almeno 50% dei casi qual è il motivo della domanda. Oppure può essere semplicemente un impiegato ben educato e voi siete un vecchio dall'aria sperduta, o una ragazza carina.

Il teorema ha una varietà di applicazioni. Nel SUPERMERCATO della vita, supponiamo che in un Paese occorrano, in base a statistiche accurate, N=100 000 dentisti operativi contemporaneamente. Supponiamo che la vita professionale di un dentista sia T=40 anni in media. Quanti dentisti devono incominciare ad esercitare ogni anno? La risposta è immediata, ne occorrono r=2500/anno. Quindi se in un certo anno si presentano 25 000 candidati al corso di laurea in odontoiatria, già sappiamo che prima o poi solo la decima parte di questi candidati farà il dentista. Lo stesso vale per ogni professione ed attività in un Paese. E aggiungo che ogni Paese civile dovrebbe essere in possesso di statistiche adeguate, ricordando che un errore si manifesta in un tempo confrontabile con T, e lo si corregge in altrettanto tempo, se ci si riesce. Ad esempio, supponendo che il sistema produca solo 1500 dentisti ogni anno, ci si accorgerà del problema quando, diciamo, ci saranno 80000 dentisti invece di 100000. Ci vorranno (100000-80000)/(2500-1500) = 20 anni per rendersi conto del problema. Per tornare al numero necessario N= 100000 se ne dovranno produrre, ad esempio, 4000 all'anno per 5 anni, facendo attenzione che questo ritmo non potrà reggere, altrimenti nel giro di 40 anni ci troveremo con 160000 dentisti invece dei 100000 che ci servono.

Una domanda interessante è se nel nostro SUPERMERCATO si guadagni qualcosa facendo un'unica coda, che permette l'accesso a diverse casse. In questo caso i clienti non sarebbero obbligati a restare bloccati in una coda in cui un cliente pasticcione sta impiegando un tempo incredibile per comprare una bottiglia d'acqua minerale, ma andrebbero sempre alla prima cassa che si libera. Dal punto di vista del "comfort" dei clienti, questo metodo è assai migliore (a meno che il luogo sia molto grande e si debba correre a casse libere e lontane come canta Johnson nella "Fanciulla del West", cosa poco gradita a persone anziane). Dal punto di vista dei numeri, invece, l’effetto sarà trascurabile, a meno che il numero di casse bloccate non provochi una diminuzione sensibile del numero di uscite con aumenti insostenibili dell’attesa nell’unica coda, il che è raro. A questo punto il manager potrà prendere diverse decisioni, che in genere equivalgono all’aggiunta (magari temporanea) di una o due nuove casse che riportano la coda in attesa a numeri tollerabili e il numero di uscite pari a quello di entrate. D’altra parte, se la coda in attesa si svuoterà rapidamente, il manager chiuderà una o due casse. In pratica la coda dovrebbe esser lunga, nel nostro caso, 5 x (numero di casse) e darebbe sempre 6 clienti in uscita, numero eguale a quello dei clienti in entrata.

Tornando alla lunghezza della coda alla CASSA nel nostro vecchio modello, essa è più o meno determinata dal manager ed in un caso semplice come quello considerato ciò è probabilmente vero. Essa è, insomma, un dato che possiamo osservare, ma non ricavare.   Esistono però dei modelli relativamente semplici che permettono di stabilire qual è la lunghezza della coda. In tal caso, questa non si forma per la decisione di una persona, ma perché i clienti non arrivano colla regolarità che abbiamo presunto. Quindi, dato un certo valor medio di frequenza di arrivo dei clienti, ogni tanto ci aspettiamo tempi morti, ed ogni tanto ci aspettiamo arrivi più frequenti. Gli arrivi più frequenti origineranno una coda.

Sono tentato di spiegare il più semplice di questi modelli. Tuttavia esso funziona solo se l’intervallo medio di arrivo fra clienti è maggiore del tempo che occorre per sbrigare un cliente.
Inoltre il modello prevede sempre dei tempi morti in cui il cassiere è ozioso. Ciò non è sempre vero, e per questa ragione ho preferito dire che la lunghezza della coda è uno dei dati del problema.