

Concetti basilari dell'analisi probabilistica

L'analisi probabilistica può essere costruita a partire da alcuni termini il cui significato è riportato qui di seguito in maniera assai semplificata, ma sufficiente a farsi una prima idea. Alla base dell'analisi probabilistica è posto il concetto di **esperimento aleatorio**, inteso come una qualsivoglia **azione** (ovvero osservazione fisica) il cui esito non possa essere stabilito a priori con certezza. Con certezza è solo possibile descrivere tutti gli **esiti** (o risultati o eventi) ammissibili dell'esperimento stesso; dunque, è naturale porsi il problema di quantificare la misura della possibilità relativa che il singolo esito o una combinazione di esiti abbiano di verificarsi. La soluzione a tale problema è appunto l'analisi probabilistica.

- **Esperimento aleatorio**

Lancio di una moneta; estrazione del numero a tombola;

Osservazione di un sistema; accesso ad una risorsa.

- **Spazio dei risultati**

Testa e croce; 1, 2, 3, ..., 88, 89, 90;

Numero di posti occupati, ..., buffer pieno / buffer vuoto;

Risorsa libera / risorsa occupata, ..., quanto si aspetta?

- **Evento**

Singolo risultato o insieme di risultati. E' una "condizione di stato" e non un "passaggio di stato", a differenza del linguaggio comune.

Ad esempio:

Numero di posti occupati o risorsa libera oppure occupata

- **Probabilità**

Quantifica la possibilità (relativa) che si realizzi un evento, tra tutti quelli individuati come possibili. Improbabile ≠ impossibile.

- **Variabile aleatoria**

Codifica uno o più eventi con numeri dell'asse reale, perché si vuole lavorare con i numeri. Introduce il concetto di durata di un evento

- **Processo stocastico**

Introduce il concetto di tempo come contenitore di eventi che hanno una certa durata.

Modella la ripetizione nel tempo di un esperimento aleatorio e, dunque, riproduce la dinamica degli stati di un sistema.

I passi dell'analisi probabilistica

1. IDENTIFICAZIONE dello "spazio" dei risultati elementari possibili;
2. ASSEGNAZIONE delle "probabilità" ai risultati possibili;
3. DEFINIZIONE degli "eventi" di interesse;
4. CALCOLO delle "probabilità di eventi" (di interesse)

dove:

- un risultato "elementare", ovvero un evento elementare, è tale perché è direttamente osservabile ed è facile assegnargli una probabilità; (lancio una moneta) (quante volte esce testa se lancio la moneta)
- un evento o risultato "d'interesse" corrisponde, in generale, ad un sottoinsieme dei risultati elementari, cioè alla combinazione logica di eventi elementari;
- l'evento di interesse deve essere posto in relazione agli eventi elementari mediante l'ALGEBRA DEGLI EVENTI (che sarà introdotta nel seguito).

La probabilità di un evento di interesse, si calcola sommando la probabilità degli eventi elementari che lo compongono.

ESEMPIO: il docente del corso si avvicina all'ultima fila di studenti in aula per osservare se "dormono più della metà". Il risultato elementare corrisponde allo stato di sonno del singolo studente, mentre l'esito "dormono più della metà" è un risultato d'interesse che non è elementare.

L'assegnamento dei valori di probabilità P agli eventi elementari deve rispettare i seguenti **assiomi**:

- (A1) per ogni evento $A \in \Omega$ (*Spazio degli eventi*) si ha $P(A) > 0$; prob non è nulla
- (A2) $P(\Omega) = 1$; per noi tutta la probabilità è pari a 1
- (A3) $P(A \cup B) = P(A) + P(B)$.

La generalizzazione di detti assiomi ad eventi non elementari è la seguente

(A1') per ogni evento $A \in \Omega$ si ha $P(A) \geq 0$;

(A2') $P(\Omega) = 1$;

(A3') $P(A \cup B) = P(A) + P(B)$ nell'ipotesi che $A \cap B = \emptyset$ (eventi disgiunti).

Si vedrà tra
molto il perché

Il principio di induzione consente di estendere l'assioma (A3') al caso dell'unione di "n" eventi disgiunti:

$$P(A_1 \cup A_2 \cup \dots \cup A_n) = \sum_{i=1}^n P(A_i).$$

Ma per trattare uno spazio di risultati numerabili all'infinito occorre sostituire l'assioma (A3') con il seguente:

$$(A3'') \quad P(\bigcup_{i=1}^{\infty} A_i) = \sum_{i=1}^{\infty} P(A_i) \leq 1$$

In tal modo, se l'unione all'infinito degli eventi A_i produce un altro evento allora il valore (compreso fra 0 e 1) a cui converge la serie delle $P(A_i)$ rappresenta la probabilità di quell'evento.

Esempio sull' unione all'infinito di eventi

Sia A_i l'evento seguente: "i-1 oggetti sono presenti in un buffer", con $i-1 = 0, 1, 2, \dots, n, \dots$

E si assegnino le seguenti:

$$P(A_i) \hat{=} (1-\rho) \cdot \rho^{i-1}, \quad : 0 < \rho < 1, \quad i = 1, 2, \dots, n, \dots$$

Allora:

$$\begin{aligned} P(\bigcup_{i=1}^{\infty} A_i) &= \sum_{i=1}^{\infty} P(A_i) \\ &= \sum_{i=1}^{\infty} (1-\rho) \rho^{i-1} = (1-\rho) \cdot \sum_{j=0}^{\infty} \rho^j \quad \text{La serie geometrica converge} \\ &= (1-\rho) \cdot \frac{1}{1-\rho} = 1 \quad \text{a } 1/(1-p) \quad \text{infatti: } \bigcup_{i=1}^{\infty} A_i \equiv \Omega \end{aligned}$$

Serie
geometrica!

8

Continua esempio sull' unione all'infinito di eventi

Si consideri adesso la seguente:

$$P(\bigcup_{i=n}^{\infty} A_i), \quad n \geq 2$$

Allora:

$$\begin{aligned} &= \sum_{i=n}^{\infty} (1-\rho) \rho^{i-1} = \sum_{j=0}^{\infty} (1-\rho) \cdot \rho^{n-1} \cdot \rho^j \quad (j \hat{=} i-n) \\ &= (1-\rho) \cdot \frac{1}{1-\rho} \cdot \rho^{n-1} = \rho^{n-1} < 1 \end{aligned}$$

$$\text{infatti: } \bigcup_{i=n}^{\infty} A_i \subset \Omega \quad \text{se } n \geq 2$$

e corrisponde all'evento: "almeno $n-1$ oggetti nel buffer"

9

Esempio sui passi dell'AP

Lo studente Giuseppe Rossi deve sostenere l'esame del corso "Analisi Probabilistica". L'esame è composto da due prove in sequenza, senza il meccanismo di ammissione alla seconda: finita la prima prova si passa sempre alla seconda. Il docente valuta le due prove insieme, dunque solo alla fine della seconda. Nella valutazione di ciascuna prova, il docente attribuisce un punteggio da 1 a 5 e lo studente, dal canto suo, **deve superare entrambe le prove con almeno 2 punti per essere promosso.** (quindi ho un vincolo)

Invece, per il prossimo appello d'esame, il docente ha deciso:

- di assegnare il punteggio di ciascuna prova (1..5) in **maniera completamente casuale;**
- di promuovere lo studente se viene fuori un punteggio cumulativo di **almeno 6** punti sull'insieme delle due prove.

Calcolare la probabilità che Rossi venga promosso:

- proprio col minimo punteggio (6); quindi calcolo le combinazioni che danno 6
- proprio col massimo punteggio (10);
- **o col minimo punteggio (6) o col massimo punteggio (10).** da notare l'OR logico

Svolgimento

Passo 1)

Identificazione dello spazio dei risultati possibili Ω

Lo spazio dei risultati Ω è dato da tutte le possibili coppie (i, j) con i il valore di punteggio ottenuto al primo compito $i = 1..5$ e j il valore di punteggio ottenuto al secondo compito $j = 1..5$. Quindi:

$$\Omega = \left\{ \begin{matrix} (1,1) & (1,2) & (1,3) & (1,4) & (1,5) \\ (2,1) & (2,2) & (2,3) & (2,4) & (2,5) \\ (3,1) & (3,2) & (3,3) & (3,4) & (3,5) \\ (4,1) & (4,2) & (4,3) & (4,4) & (4,5) \\ (5,1) & (5,2) & (5,3) & (5,4) & (5,5) \end{matrix} \right\}.$$

Passo 2)

Assegnazione delle probabilità ai risultati possibili

Poiché il docente ha deciso di assegnare il punteggio di ciascuna prova in maniera completamente casuale, ogni evento elementare (i, j) è equiprobabile per cui

$$P(i, j) = \frac{1}{25} \quad \forall (i, j).$$

Passo 3) Definizione degli eventi di interesse

Stabilito che la promozione si ottiene conseguendo un punteggio complessivo sulle due prove di valore pari ad almeno 6, gli eventi di interesse si riferiscono alla promozione dello studente Rossi proprio col minimo punteggio (6), proprio col massimo punteggio (10) o col minimo punteggio (6) o col massimo punteggio (10).

Risoluzione) Calcolo delle probabilità di eventi di interesse

Definiamo i seguenti eventi:

A $\hat{=}$ promozione con un valore di punteggio pari proprio al minimo (6)

B $\hat{=}$ promozione con un valore di punteggio pari proprio al massimo (10)

C $\hat{=}$ promozione con un valore di punteggio pari al minimo (6) o al massimo (10)

allora

$$P(A) = \frac{3}{25} = 0,12 \text{ (i.e. 3 casi su 25 } \{(2,4), (3,3), (4,2)\}\text{)}$$

$$P(B) = \frac{1}{25} = 0,04 \text{ (i.e. 1 caso su 25 } \{(5,5)\}\text{)}$$

$$P(C) = P(A) + P(B) = \frac{4}{25} = 0,16.$$

Evento limite e sua probabilità

Data una sequenza di eventi $\{E_n, n \geq 1\}$ si dirà che essa è una **sequenza crescente** se risulta $E_1 \subset E_2 \subset \dots \subset E_n \subset E_{n+1} \subset \dots$. Se, invece, risulta $E_1 \supset E_2 \supset \dots \supset E_n \supset E_{n+1} \supset \dots$ allora si dirà che la **sequenza è decrescente**.

E1 sarebbe superset ovvero insieme che contiene E2.

Facendo riferimento a una sequenza crescente si definisce **evento limite** e si indica con $\supset E$ il seguente:

$$\supset E \hat{=} \lim_{n \rightarrow \infty} E_n = \lim_{n \rightarrow \infty} \bigcap_{i=1}^n E_i$$

Viceversa, facendo riferimento a una sequenza decrescente si definisce analogamente:

$$\subset E \hat{=} \lim_{n \rightarrow \infty} E_n = \lim_{n \rightarrow \infty} \bigcup_{i=1}^n E_i$$

Proposizione:

Se $\{E_n, n \geq 1\}$ è una sequenza di eventi crescente o decrescente e si indica semplicemente con E l'evento limite, allora risulta:

$$P(E) \hat{=} P\left(\lim_{n \rightarrow \infty} E_n\right) = \lim_{n \rightarrow \infty} P(E_n)$$

si vedrà in futuro cosa vuol dire

PROVA (per una sequenza crescente):

$$\begin{aligned} P(E) &\hat{=} P\left(\lim_{n \rightarrow \infty} E_n\right) = P\left(\lim_{n \rightarrow \infty} \bigcup_{i=1}^{n-1} E_i\right) = P\left(\bigcup_{i=1}^{\infty} E_i\right) = P\left(E_1 \cup \bigcup_{i=1}^{\infty} E_{i+1} \setminus E_i\right) \\ &= P(E_1) + \sum_{i=1}^{\infty} [P(E_{i+1}) - P(E_i)] = P(E_1) + \lim_{n \rightarrow \infty} \sum_{i=1}^n [P(E_{i+1}) - P(E_i)] = \lim_{n \rightarrow \infty} P(E_n) \end{aligned}$$

La proposizione formalizza un concetto di continuità dell'assegnamento probabilistico che sarà utile in seguito, quando si lavorerà con funzioni continue che effettuano l'assegnamento probabilistico sulla semiretta reale e si porrà il problema di attribuire una probabilità tanto ad un intervallo di reali quanto ad un unico punto. Allora, l'intervallo e il punto saranno, rispettivamente, l'evento e l'evento limite.

Esempio su evento limite e sua probabilità

Sia E_i l'evento seguente:

“ al più i oggetti sono presenti nel buffer”,
con $i = 1, 2, 3, \dots$ e quindi E_1, E_2, \dots una sequenza crescente.

Poiché:

$$P(E_i) = 1 - \rho^i, \quad : 0 < \rho < 1, \quad i = 1, 2, \dots \quad \text{come visto in precedenza}$$

Allora:

$$\begin{aligned} P(E) &= P(\bigcup_{i=1}^{\infty} E_i) = P(E_1) + \sum_{i=1}^{\infty} [P(E_{i+1}) - P(E_i)] \\ &= (1 - \rho) + \sum_{i=1}^{\infty} [(1 - \rho^{i+1}) - (1 - \rho^i)] \\ &= \lim_{n \rightarrow \infty} P(E_n) = \lim_{n \rightarrow \infty} (1 - \rho^n) = 1 \end{aligned}$$

11

Costruzione di un modello di assegnazione

$$P : [0,1,2,\dots,n] \rightarrow [0,1] \in \Re$$

$$\sum_{n=0}^{\infty} \frac{x^n}{n!} = e^x, \quad x > 0 \quad (\text{MODELLO DI PARTENZA})$$

$$\sum_{n=0}^{\infty} \frac{x^n}{n!} e^{-x} = 1, \quad x > 0 \quad (\text{PASSAGGIO INTERMEDI})$$

$$0 \leq \frac{x^n}{n!} e^{-x} \leq 1, \quad n = 0, 1, 2, \dots \quad (\text{VERIFICA FINALE})$$

$P(N = n) = \frac{x^n}{n!} e^{-x}, \quad n = 0, 1, 2, \dots$	MODELLO DI POISSON
--	---------------------------

12

Il modello di Poisson è un modello che nasce dalla teoria che vedremo molto più in la nel corso. Nel mondo reale si utilizza per le chiamate in un call center, cioè o la chiamata ha vita breve e dura pochissimi secondi, oppure la chiamata dura un tempo ragionevolmente lungo (non si sta in chiamata per sempre, cioè all'infinito)

11

Problemi combinatori

Quando lo spazio dei risultati è costituito da un numero finito di eventi elementari equiprobabili, la probabilità di un "tipo di risultato" (evento di interesse) si calcola come rapporto tra il numero di risultati elementari che definiscono quel evento d'interesse e il numero di risultati (eventi) elementari che definiscono tutto lo spazio dei risultati.

Calcolare la probabilità di un evento d'interesse è un problema combinatorio (beccare il 6 al Superenalotto, ad esempio) che può essere risolto a partire dai seguenti calcoli, basilari, di numerosità:

- La numerosità del campione ordinato (disposizione) di "k" risultati estratti da uno spazio di "n", col meccanismo del rimpiazzo dell'oggetto estratto è pari a:

$$\underbrace{n \cdot n \cdots n}_{k \text{ volte}} = n^k \quad \begin{array}{l} \text{esperimento con} \\ \text{risultati ripetibili} \end{array}$$

Esempio

Si consideri l'esperimento "lancio di una moneta" per il quale il possibile spazio dei risultati è $\Omega = \{T, C\}$ ($n = 2$). Posto che il rimpiazzo è automatico per la natura dell'esperimento, se il lancio della moneta viene effettuato tre volte ($k = 3$), allora le possibili disposizioni sono:

T, T, T

T, T, C

T, C, T

C, T, T ovvero $n^k = 2^3 = 8$

C, C, T

C, T, C

T, C, C

C, C, C

- La numerosità del campione ordinato (disposizione) di "k" risultati estratti da uno spazio di "n", senza rimpiazzo è pari a:

$$n(n-1)(n-2) \cdots (n-(k-1)) = \frac{n!}{(n-k)!} \quad \begin{array}{l} \text{esperimento} \\ \text{risultati non ripetibili} \end{array}$$

Fissando $k \triangleq n$ (permutazione) si ottiene: $n!$ quindi estraggo tutti gli elementi
Esempio

Per ricoprire 2 posizioni in un'azienda ci sono 6 concorrenti. Dal momento che tutti i concorrenti hanno lo stesso bagaglio di esperienze, si decide di estrarre a caso i due nominativi dei futuri promossi. Ovviamente se un concorrente viene estratto, non può essere candidato anche per l'altra posizione. Quindi, le possibili coppie estraibili (disposizioni senza rimpiazzo) sono:

$$\frac{n!}{(n-k)!} = \frac{6!}{(6-2)!} = 30 \quad \begin{array}{l} \text{se fossero 6 posizioni allora} \\ \text{avrei 720 permutazioni possibili} \end{array}$$

- Campione non ordinato (combinazione) di "k" risultati estratti da uno spazio di "n", senza il meccanismo del rimpiazzo.

$$\frac{n!}{k!(n-k)!} \triangleq \binom{n}{k}.$$

Nel Superenalotto la combinazione vincente del "6" si ottiene dal primo numero estratto su 6 ruote diverse. Il meccanismo di estrazione garantisce che uno stesso numero non possa essere estratto come primo numero su più di una delle 6 ruote. In altre parole, si tratta di un meccanismo di estrazione casuale senza rimpiazzo.

Dunque, si possono estrarre 90 numeri sulla prima ruota, 89 sulla seconda, 88 sulla terza, 87 sulla quarta, 86 sulla quinta e 85 sulla sesta, ovvero:

$$n \cdot (n-1) \cdot (n-2) \cdot (n-3) \cdot (n-4) \cdot (n-(k+1)) = \frac{n!}{(n-k)!}$$

Se si pensa alla sestina vincente, l'ordine con cui vengono estratti i numeri non è importante, quindi, quella sestina può uscire in $k!$ modi diversi, ovvero la numerosità si riduce di $k!$ risultando:

$$\frac{n!}{k!(n-k)!} = \binom{90}{6}.$$

mentre la probabilità aumenta di $k!$, visto che è pari al reciproco della numerosità.

Il precedente è noto come *coefficiente binomiale* (e si legge n su k).

Ipotizzando di avere 1024 thread resi disponibili dalla CPU, separati per tipologie ad esempio thread I/O, thread di rete, thread di calcolo e thread di scrittura. Supponiamo che un processo per essere completato debba avere un thread di ogni tipo, perciò 4. Allora si ha che un processo è completato con coefficiente binomiale 1024 su 4

Il problema dei compleanni

E' dato un gruppo di «k» persone aggregate a caso e si vuole calcolare la probabilità che almeno due di esse siano nate nello stesso giorno di anni sia pur differenti.

Soluzione:

(calcoliamo la probabilità dell'evento d'interesse come il complemento ad 1 della probabilità dell'evento complementare a quello d'interesse)

$$= 1 - \frac{n(n-1)(n-2) \cdots (n-(k-1))}{n^k} \quad \text{con } n = 365$$

In particolare,

con k=23 risulta = 0,5073 e con k=50 risulta = 0,9704

4

Quesito finale

Quanto vale la probabilità (P) che uno qualsivoglia di «n» risultati (elementari ed equiprobabili) si realizzi proprio «h» volte, in corrispondenza della ripetizione dell'esperimento aleatorio «k» volte?

(ovviamente deve essere $k \geq h$)

$$P = \frac{\binom{k}{h} (n-1)^{k-h}}{n^k} \quad h = 0, 1, \dots, k$$

5

Piccolo schema per ricordarsi:

- campione ordinato -> disposizione (con rimpiazzo: esponenziale, senza: rapporto di fattoriali)
- campione non ordinato -> combinazione (coefficiente binomiale)

Una applicazione del calcolo combinatorio

Sia S un insieme di n elementi, di cui d difettosi ($d \leq n$). Si vuole calcolare la probabilità di ottenere r elementi difettosi, estraendo k ($k \leq n$) elementi a caso e senza rimpiazzo dall'insieme S .

Il numero di possibili campioni casuali di k elementi, estratti da un insieme di n senza il rimpiazzo è:

$$\binom{n}{k} \quad \text{appunto coefficiente binomiale}$$

Analogamente, i campioni di r elementi difettosi che possono capitare, dai d elementi difettosi di S , ammontano a:

$$\binom{d}{r}$$

Infine, il numero di campioni non difettosi di $k-r$ elementi, appartenenti al campione dei k elementi estratti ammonta a:

$$\binom{n-d}{k-r}$$

(Si noti che i $k-r$ elementi appartengono all'insieme degli $n-d$ elementi non difettosi in S .)

Poiché gli eventi elementari sono equiprobabili, è possibile calcolare la probabilità di ottenere r elementi difettosi dai k elementi estratti come:

prodotto tra elementi difettosi e non difettosi in rapporto ad elementi generici $\longrightarrow P(r | k, d, n) = \frac{\binom{d}{r} \binom{n-d}{k-r}}{\binom{n}{k}}$ Ipergeometrica!

La precedente probabilità è nota come *probabilità Ipergeometrica*. Un'ulteriore considerazione riguarda il numero r di elementi difettosi dal campione estratto di k elementi:

$$\max\{0, d + k - n\} \leq r \leq \min\{d, k\}.$$

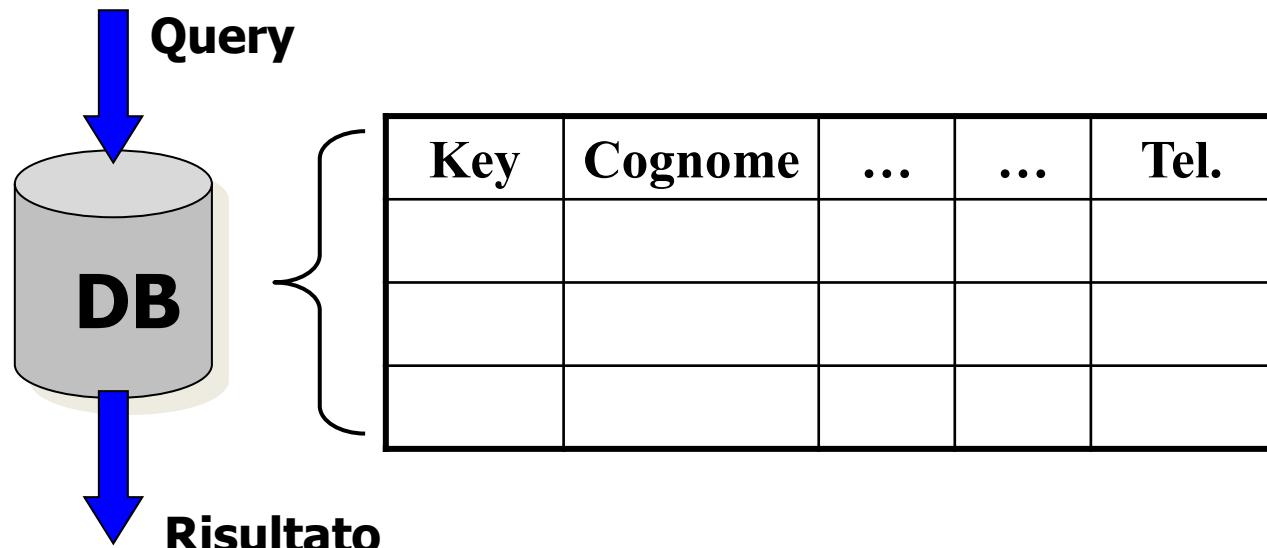
Esempio sulla probabilità ipergeometrica

Una tabella di database contiene informazioni di 10 persone, tra cui i rispettivi numeri di telefono. Però 4 numeri sono inconsistenti.

Se viene estratto un campione di 6 numeri telefonici, qual è la probabilità che 2 dei 6 numeri telefonici estratti risultino inconsistenti?

RISPOSTA:

$$\begin{aligned} P(2|6,4,10) \\ = \binom{4}{2} \cdot \binom{6}{4} / \binom{10}{6} \\ = 0.43 \end{aligned}$$



Algebra degli eventi

È completamente definita dalle seguenti 5 leggi:

Legge Comutativa: $A \cup B = B \cup A$ $A \cap B = B \cap A$

Legge Associativa: $A \cup (B \cup C) = (A \cup B) \cup C$ $A \cap (B \cap C) = (A \cap B) \cap C$

Legge Distributiva: $A \cup (B \cap C) = (A \cup B) \cap (A \cup C)$ $A \cap (B \cup C) = (A \cap B) \cup (A \cap C)$

Legge dell'Identità: $A \cup \emptyset = A$ $A \cap \Omega = A$

Legge del Complemento: $A \cup \bar{A} = \Omega$ $A \cap \bar{A} = \emptyset$

Ogni relazione fra eventi di interesse ed eventi elementari può essere stabilita usando opportunamente le leggi appena esposte.

∩ intersezione negli insiemi, e (congiunzione appunto probabilità congiunta) nelle probabilità

Relazioni valide:

1. $A \cup A = A$; $A \cap A = A$; $A \cup \Omega = \Omega$; $A \cap \emptyset = \emptyset$;
2. $A \cap (A \cup B) = A$; $A \cup (A \cap B) = A$;
3. $\overline{A \cup B} = \bar{A} \cap \bar{B}$; $\overline{A \cap B} = \bar{A} \cup \bar{B}$;
4. $\overline{\overline{A}} = A$; $A \cup (\bar{A} \cap B) = A \cup B$.

A titolo di esempio:

$$(A \cap B) \cup (\bar{A} \cap B) = B$$

$$(A \cap B) \cup (\bar{A} \cap B)$$

$$(B \cap A) \cup (B \cap \bar{A}) \quad \text{(legge commutativa)}$$

$$B \cap (A \cup \bar{A}) \quad \text{(legge distributiva)}$$

$$B \cap \Omega \quad \text{(legge del complemento)}$$

$$B \quad \text{(legge dell'identità)}$$

Vedremo nelle prossime pagine come:

$$P(A \cap B) = P(A) \cdot P(B) \quad \text{se e solo se } A \text{ e } B \text{ sono indipendenti}$$

$$P(A \cup B) = P(A) + P(B) \quad \text{se } A \text{ e } B \text{ sono eventi disgiunti (intersezione nulla, NON POSSONO accadere contemporaneamente!)}$$

Esempio evento congiunto: da un sacco di palline colorate e numerate estraggo una pallina rossa e pari (evento congiunto perché può accadere contemporaneamente)

Arrivano datagrammi dal protocollo IP, il datagramma che prende il sistema è quello id 1 e destinazione proprio il sistema.

Diagrammi di Venn

Un tipo di rappresentazione grafica degli eventi, molto utile per illustrare le leggi che definiscono l'algebra degli eventi, sono i diagrammi di Venn. Lo spazio degli esiti Ω viene rappresentato da un grande rettangolo. Gli eventi da prendere in considerazione, invece, sono rappresentati da cerchi o altre curve chiuse disegnate all'interno del rettangolo. A questo punto, tutti gli eventi non elementari di interesse possono essere evidenziati colorando opportune regioni del diagramma. Ad esempio nei tre diagrammi di Venn illustrati nella figura a seguire, le regioni scurite rappresentano, nell'ordine, gli eventi $A \cup B$, $A \cap B$ ed \bar{A} . Il diagramma di Venn della seconda figura, invece, mostra che $B \subset A$.

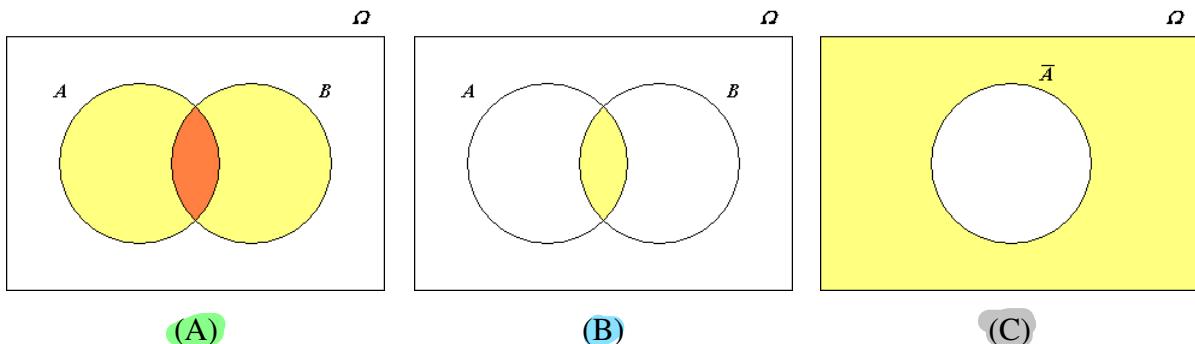


Figura – Diagrammi di Venn per l'unione, l'intersezione e il complemento

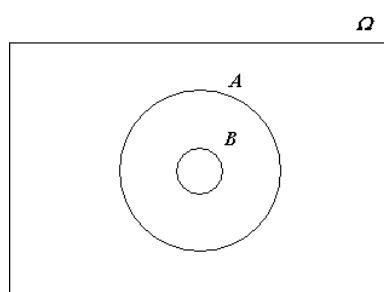


Figura – Diagramma di Venn che illustra la relazione $B \subset A$

Nella terza figura i diagrammi di Venn sono impiegati per verificare la seconda delle due proprietà distributive: $(C \cap A) \cup (C \cap B) = C \cap (A \cup B)$.

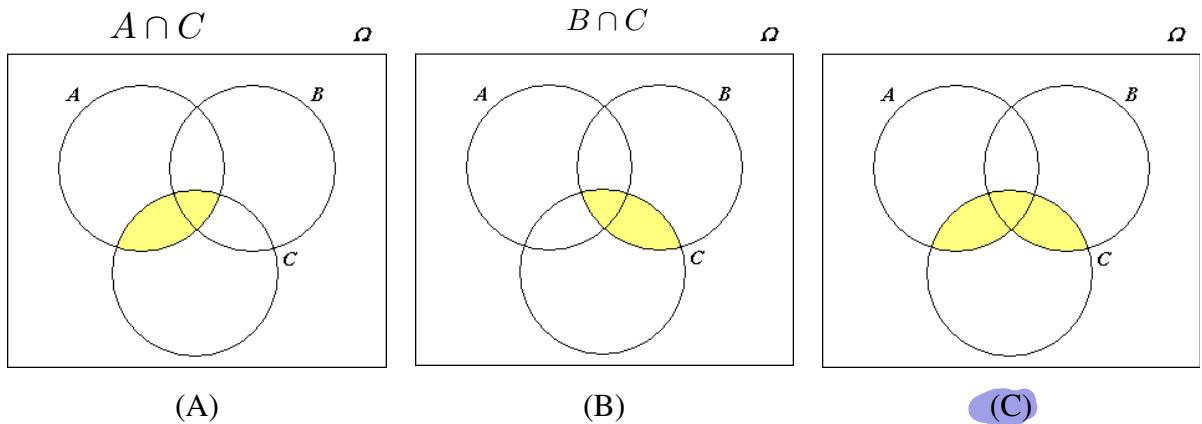


Figura – Illustrazione di una proprietà distributiva per mezzo dei diagrammi di Venn

Per esercizio, si propone di verificare mediante l'uso dei diagrammi di Venn le seguenti relazioni note come **Leggi di De Morgan**:

$$\overline{(A \cup B)} = \overline{A} \cap \overline{B} \quad \text{risultato tutto lo spazio del rettangolo esclusi } A \text{ e } B$$

$$\overline{(A \cap B)} = \overline{A} \cup \overline{B} \quad \text{risultato solo l'intersezione tra } A \text{ e } B$$

DEFINIZIONI

- Gli eventi A_1, \dots, A_n sono detti **mutuamente esclusivi** se e solo se risulta qualsiasi coppia ha intersezione nulla $A_i \cap A_j = \begin{cases} A_i & \text{se } i = j \\ \emptyset & \text{altrimenti} \end{cases} \quad (i = 1 \dots n, j = 1 \dots n)$
- Gli eventi A_1, \dots, A_n sono detti **collettivamente esaustivi** se e solo se risulta: $A_1 \cup A_2 \cup \dots \cup A_n = \Omega$
- Gli eventi A_1, \dots, A_n costituiscono una **partizione di Ω** se e solo se risultano mutuamente esclusivi e collettivamente esaustivi.

Probabilità dell'unione di 2 eventi:

Di seguito l'unione della probabilità tra 2 eventi ed il suo caso generale con n elementi

$$P(A \cup B) = P(A) + P(B) - P(A \cap B)$$

Probabilità dell'unione di eventi

$$P(A \cup B) = P(A) + P(B) - P(A \cap B)$$

(DUE EVENTI)

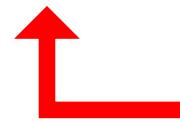
Prova:

Sia $A \cup B = A \cup (\bar{A} \cap B)$ e $B = (A \cap B) \cup (\bar{A} \cap B)$ sostituendo P(B)

$$P(A \cup B) = P(A) + P(\bar{A} \cap B) \quad P(B) = P(A \cap B) + P(\bar{A} \cap B)$$

da cui ottengo che

$$P(A \cap B) \geq P(A) + P(B) - 1$$



$$P(\bar{A} \cap B) = P(B) - P(A \cap B)$$

Disuguglianaza di Bonferroni

La probabilità dell'unione di n eventi è minore uguale alla sommatoria delle probabilità di n elementi

(n EVENTI)

$$P\left(\bigcup_{i=1}^n A_i\right) = P(A_1 \cup A_2 \cup \dots \cup A_n) =$$

probabilità dell'unione di N eventi
NON disgiunti

$$= \sum_{\forall i} P(A_i) - \sum_{1 \leq i < j \leq n} P(A_i \cap A_j) +$$

$$+ \sum_{1 \leq i < j < k \leq n} P(A_i \cap A_j \cap A_k) + \dots + (-1)^{n-1} P(A_1 \cap A_2 \cap \dots \cap A_n)$$

Disuguaglianza di Boole

$$P\left(\bigcup_{i=1}^n A_i\right) \leq \sum_{i=1}^n P(A_i)$$

Disuguaglianze di Bonferroni

Nel caso di collezioni finite di eventi, la precedente disuguagliazione può venire generalizzata nelle cosiddette **disuguaglianze di Bonferroni** le quali forniscono estremi superiori e inferiori alla probabilità per l'unione di tali eventi.

Introduciamo le seguenti quantità:

$$S_1 := \sum_{i=1}^n P(A_i) ,$$

$$S_2 := \sum_{i < j} P(A_i \cap A_j) ,$$

e per $2 < k \leq n$,

$$S_k := \sum P(A_{i_1} \cap \cdots \cap A_{i_k}) ,$$

dove si intende che la sommazione sia da effettuare sopra tutte le k -uple di interi i_1, i_2, \dots, i_k soddisfacenti $i_1 < i_2 < \cdots < i_k$.

Per gli interi dispari $k \geq 1$ si dimostra che

$$P\left(\bigcup_{i=1}^n A_i\right) \leq \sum_{j=1}^k (-1)^{j+1} S_j ,$$

mentre per gli interi pari $k \geq 2$

$$P\left(\bigcup_{i=1}^n A_i\right) \geq \sum_{j=1}^k (-1)^{j+1} S_j .$$

Probabilità condizionata

Si supponga che A sia un evento di interesse (del quale, cioè, si vuole calcolare la probabilità, $P(A)$) e che esso abbia intersezione non vuota con un altro, B, che ha probabilità, $P(B)$, già nota e non nulla. Dunque $A \cap B \neq \emptyset$

La probabilità condizionata $P(A|B)$ misura la probabilità che, come risultato dell'esperimento aleatorio, si realizzi (anche) l'evento A, sotto la condizione che si realizzerà B con certezza. Nel senso che si pone come ipotesi l'esito seguente dell'esperimento: il risultato è uno degli eventi elementari (ω) la cui unione definisce B e vale $P(B)$.

Si capisce che A ha una sorta di dipendenza da B. B accade ed influenza A!

Definizione:

$$P(A|B) \triangleq \sum_{\omega \in A \cap B} P(\omega | B \text{ certo}) = ?$$

Da B certo $\implies P(B) = 1$

Si pone:

$$P(\omega | B \text{ certo}) \equiv P(\omega) = \frac{P(\omega)}{P(B)} \quad \forall \omega \in B$$

$$P(\omega | B \text{ certo}) \equiv P(\omega) = 0 \quad \forall \omega \in \overline{B} \cap \Omega$$

E banalmente si verifica che:

$$\sum_{\omega \in B} \hat{P}(\omega) = \sum_{\omega \in B} \frac{P(\omega)}{P(B)} = 1$$

Dunque:

$$P(\omega | B \text{ certo}) = \frac{P(\omega)}{P(B)} \quad \omega \in A \cap B$$

e si ottiene:

$$P(A|B) = \sum_{\omega \in A \cap B} \frac{P(\omega)}{P(B)} = \frac{P(A \cap B)}{P(B)}$$

la condizionata è pari al rapporto tra probabilità congiunta e probabilità condizionante

$$P(A|B) = \frac{P(A \cap B)}{P(B)} \implies P(A \cap B) = P(A|B) \cdot P(B)$$

utile nelle manipolazioni

quindi la congiunta è pari alla condizionata per la condizionante! (dimostrazione nelle prossime pagine)

PROBABILITA' CONGIUNTA E "DIPENDENZA"

CONGIUNTA

$$P(A_1 \cap A_2 \cap A_3) = ?$$

Cattura la dipendenza
fra gli eventi.

Basta porre: $C \hat{=} A_2 \cap A_3$

$$P(A_1 \cap A_2 \cap A_3) = P(A_1 \cap C) = P(A_1 | C) \cdot P(C)$$

$$= P(A_1 | C) \cdot P(A_2 \cap A_3) \stackrel{\text{esplicito nuovamente}}{=} P(A_1 | C) \cdot P(A_2 | A_3) \cdot P(A_3)$$

$$= P(A_1 | A_2 \cap A_3) \cdot P(A_2 | A_3) \cdot P(A_3)$$

Assumendo: $P(A_1 | A_2 \cap A_3) \hat{=} P(A_1 | A_2)$

$$= P(A_1 | A_2) \cdot P(A_2 | A_3) \cdot P(A_3)$$

Quindi A1 dipende da A2 che dipende a sua volta da A3

dipendenza a catena

Dipendenza fra eventi generalizzata

È illustrata dalla formula seguente:

$$P(A_1 \cap A_2 \cap \dots \cap A_n) =$$

$$\text{suggerimento: } B \hat{=} A_2 \cap \dots \cap A_n$$

$$\text{e poi } = P[A_1 \cap B] = P[A_1 | B]P[B]$$

Formula finale
detta Chain Rule: $P(A_1 \cap A_2 \cap \dots \cap A_n) = P[A_1 | (A_2 \cap A_3 \cap \dots \cap A_n)] * \dots * P[A_{n-1} | A_n] * P[A_n]$

• Dipendenza a catena generalizzata

È definita nel seguente modo

$$\begin{aligned} P(A_1 \cap A_2 \cap \dots \cap A_n) &= P(A_n \cap A_{n-1} \cap \dots \cap A_1) \hat{=} \\ &\hat{=} P[A_n | A_{n-1}] * P[A_{n-1} | A_{n-2}] * \\ &\dots * P[A_2 | A_1] * P[A_1] \end{aligned}$$

pensare agli indici 1,...,n come una sequenza di istanti temporali ai quali sono agganciati/riferiti gli eventi corrispondenti, A1,...,An

• Indipendenza stocastica di una coppia di eventi

S'immagini di aver calcolato la probabilità di due eventi qualsivoglia ma non disgiunti (disgiunti significa logicamente incompatibili), $A \subset \Omega$ e $B \subset \Omega$, e di averla calcolata sulla base dell'assegnamento originario di probabilità ad eventi elementari di Ω . Siano esse $P(A)$ e $P(B)$, entrambe > 0 . Ciò premesso, A e B potrebbero essere definiti eventi stocasticamente indipendenti (indipendenti in probabilità) quando la modifica della probabilità già calcolata per uno dei due non induce modifiche alla probabilità calcolata per l'altro evento. La modifica di interesse pratico è quella che consiste nell'aggiornare una delle due probabilità $[P(A) \text{ o } P(B)]$ portandola al valore unitario, ovvero ipotizzare che si verifichi con certezza uno dei due eventi $[A \text{ o } B]$. Quindi, ci si chiede se la nuova $P(A|B)$ [ovvero $P(B|A)$] risulterà uguale alla vecchia $P(A)$ [ovvero $P(B)$] oppure no. In caso affermativo si potrà parlare di indipendenza stocastica. D'altra parte, tenendo presente che - grazie alla formula della probabilità congiunta - i risultati $P(A|B)=P(A)$ e $P(B|A)=P(B)$ sono entrambi implicati dall'assumere che risulti $P(A \cap B) = P(A)P(B) = P(B \cap A)$ si conviene di proporre la seguente definizione di indipendenza stocastica per la coppia di eventi A e $B \subset \Omega$:

Gli eventi A e $B \subset \Omega$ e non incompatibili, né di probabilità individuale nulla, sono indipendenti se e solo se risulta: $P(A \cap B) = P(A) \cdot P(B) > 0$

dimostrazione

$$P(A \cap B) = P(A|B)P(B) = P(B|A)P(A) = P(A)P(B) = P(B)P(A) > 0$$

Dunque la probabilità congiunta è pari al prodotto delle singole probabilità se e solo se c'è l'ipotesi di indipendenza!

E' interessante osservare che la formula (+) può essere letta alla seguente maniera:

$$\frac{P(A \cap B)}{P(A)} = \frac{P(B)}{P(\Omega)} \quad \text{ovvero} \quad \frac{P(B \cap A)}{P(B)} = \frac{P(A)}{P(\Omega)}$$

Nell'illustrazione grafica proposta, s'immagini per semplicità che l'area dello spazio Ω valga 1, come la probabilità ad esso associata:

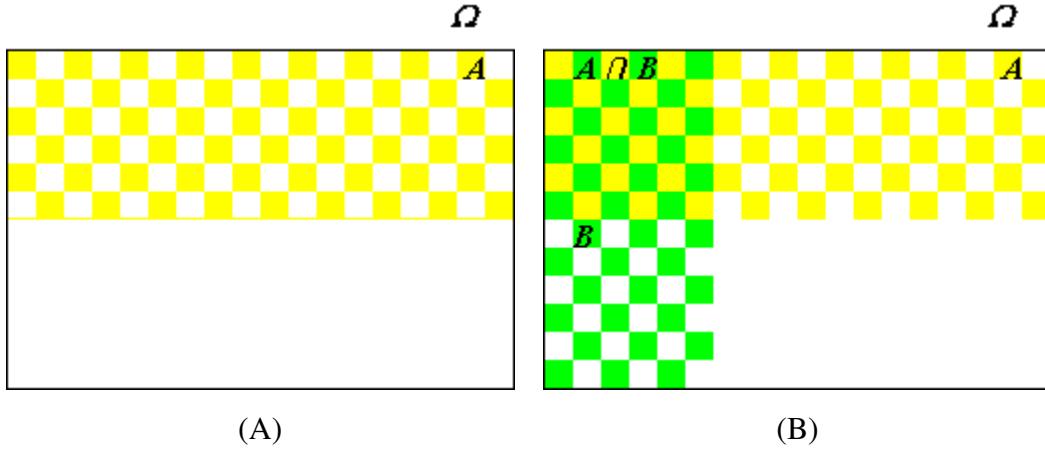


Figura – Illustrazione grafica della proprietà di indipendenza

Allora, visto che l'area di A appare proprio come la metà di quella di Ω e quella di B come $1/3$, risulta :

$$P(A) = \frac{1}{2}, \quad P(B) = \frac{1}{3}, \quad \text{ma} \quad P(A \cap B) = \frac{1}{2}$$

e quindi:

$$\frac{P(A \cap B)}{P(A)} = \frac{P(B)}{P(\Omega)} \quad \text{ovvero} \quad \frac{1/6}{1/2} = \frac{1/3}{1}$$

Dunque a e B sono stocasticamente indipendenti.

L'illustrazione grafica potrebbe servire a convincersi che il concetto di indipendenza stocastica è una proprietà che scaturisce dalla specifica rappresentazione topologica dei due eventi nello spazio campionario. Infatti, ruotando B di 90 gradi in senso orario esso diventa logicamente incompatibile con A, altro che indipendente!
Esempio

Dato un mazzo di carte da poker di 52 carte a 4 semi di cui due rossi (i.e. cuori e quadri) e due neri (i.e. fiori e picche), si considerino i seguenti eventi

$A \hat{=} \text{estrazione di un asso}; \quad B \hat{=} \text{estrazione di una carta di cuori}.$

La probabilità di questi eventi è data rispettivamente da

$$P(A) = \frac{4}{52} = \frac{1}{13}, \quad P(B) = \frac{13}{52} = \frac{1}{4}.$$

D'altra parte la probabilità di estrarre l'asso di cuori, ovvero $P(A \cap B)$ è $1/52$, come ogni altra carta, perché il mazzo non è truccato. E questo permette di verificare la validità della relazione:

$$\frac{P(A \cap B)}{P(A)} = \frac{P(B)}{P(\Omega)}. \quad \text{Infatti, risulta: } \frac{1/52}{1/13} = \frac{1/4}{1}.$$

Dunque abbiamo provato che è vera la seguente: $P(A \cap B) = P(A) * P(B)$, ovvero che A e B sono eventi stocasticamente indipendenti, senza invocare la probabilità condizionata!

Viceversa, è sufficiente trasformare B in evento certo e invocare la probabilità condizionata per ristabilire l'indipendenza stocastica fra A e B.

Infatti, si ha: $\Omega \rightarrow B$ e quindi: $P(A \cap B | B) \equiv P(A | B) = 1/13$, perché c'è solo un asso di cuori fra le 13 del seme cuori. A questo punto, il semplice confronto del valore appena ottenuto per $P(A|B)$ con $P(A)=1/13$ completa la prova.

Il "motivo" dell'indipendenza è che il rapporto tra il numero di eventi elementari che compongono B (13) e il totale degli eventi elementari (52) che compongono Ω è uguale al rapporto tra il numero di eventi elementari (1) che sono "asso \cap di cuori" ($A \cap B$) e il numero di eventi elementari che compongono A (4). Sempre $1/4$.

Si supponga adesso di lavorare con un mazzo di carte speciale, che contiene sempre 13 carte per seme ma che abbia 2 assi di cuori (buttando via il 2 di cuori, ad esempio). In tal caso, lo studente può verificare che cade l'indipendenza stocastica tra gli eventi A e B.

Il risultato sarebbe $\frac{2}{5} = \frac{1}{4}$

Indipendenza stocastica fra 3 eventi

Dati gli eventi A, B e C devono valere le seguenti condizioni di indipendenza mutua e indipendenza a coppie:

$$P(A \cap B \cap C) = P(A) \cdot P(B) \cdot P(C)$$

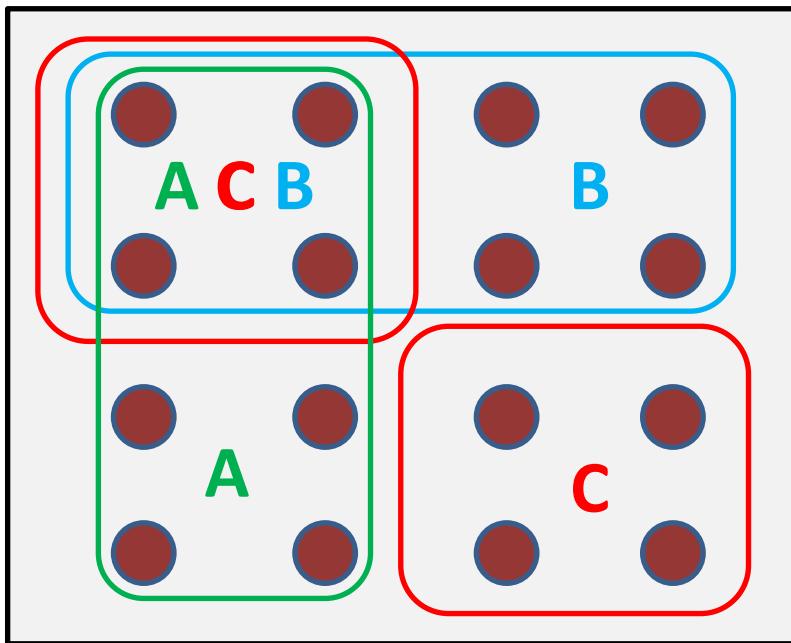
+

$$\left\{ \begin{array}{l} P(A \cap B) = P(A) \cdot P(B) \\ P(A \cap C) = P(A) \cdot P(C) \\ P(B \cap C) = P(B) \cdot P(C) \end{array} \right.$$

Se non valgono tutte, l'indipendenza stocastica non sussiste!

Generalizzare al caso di n eventi.

Pairwise independence and mutual independence



$$P(X \subseteq \Omega) \hat{=} \frac{\#\text{risultati} \in X}{\#\text{risultati} \in \Omega}$$

↓

$$\begin{aligned} P(A) &= P(B) = P(C) \\ &= 8/16 = 1/2 \end{aligned}$$

Pairwise independence and mutual independence

P $P(A \cap B) = P(A | B) \cdot P(B) = P(A) \cdot P(B)$
A $= (1/2) \cdot (1/2) = 1/4$
I

R $P(A \cap C) = P(A | C) \cdot P(C) = P(A) \cdot P(C)$
W $= (1/2) \cdot (1/2) = 1/4$
I

S $P(B \cap C) = P(B | C) \cdot P(C) = P(B) \cdot P(C)$
E $= (1/2) \cdot (1/2) = 1/4$

NO MUTUAL:

$$\begin{aligned}P(A \cap B \cap C) &= P(A | B \cap C) \cdot P(B \cap C) \\&= P(A | B \cap C) \cdot P(B | C) \cdot P(C) = 1 \cdot (1/2) \cdot (1/2) = 1/4 \\&\neq P(A) \cdot P(B) \cdot P(C) = (1/2) \cdot (1/2) \cdot (1/2) = 1/8\end{aligned}$$

Formula della probabilità totale e corollario di Bayes

Data una partizione $\{B_1, B_2, \dots, B_n\}$ dello spazio Ω e un evento di interesse (A):

$$A = A \cap \Omega = A \cap \bigcup_{i=1}^n B_i = \bigcup_{i=1}^n (A \cap B_i)$$

e da qui

$$P(A) = \sum_{i=1}^n P(A \cap B_i) = \sum_{i=1}^n P(A | B_i)P(B_i)$$

grazie all'ipotesi di partizionamento di Ω , che rende disgiunti gli eventi:

$$A \cap B_i, i = 1, \dots, n$$

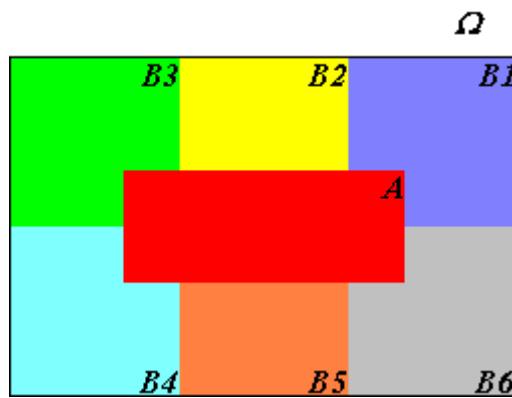


Figura – Illustrazione della formula della probabilità totale per A e $\{B_1, B_2, \dots, B_6\}$

Da notare come la totale sia una SOMMA di tutte le congiunte!

Corollario di Bayes: Importante

Scelto a caso l'evento B_j della partizione $\{B_1, B_2, \dots, B_j, \dots, B_n\}$ si ricava la formula

seguente, nota come regola di Bayes:

$$\begin{aligned} P(B_j | A) &= \frac{P(B_j \cap A)}{P(A)} = \frac{P(A \cap B_j)}{P(A)} \\ &= \frac{P(A | B_j)P(B_j)}{\sum_{i=1}^n P(A | B_i)P(B_i)} \end{aligned}$$

dove $P(B_j)$ è detta prior probability di B_j
 $P(B_j|A)$ è detta posterior probability di B_j

Essa consente di calcolare la probabilità che si realizzi un evento specifico della partizione, posto che si è realizzato l'evento di interesse.

Esempio sulla probabilità totale e sulla regola di Bayes

Consideriamo tre macchine (M_1 , M_2 , M_3) che forniscono pezzi che dovranno essere assemblati dopo un controllo di qualità:

- I pezzi difettosi, espressi in percentuale, prodotti da ogni macchina sono:
 - M_1 ----- 1% (un pezzo su cento lavorati dalla macchina M_1 è difettoso)
 - M_2 ----- 5% (cinque pezzi su cento lavorati dalla macchina M_2 sono difettosi)
 - M_3 ----- 2% (due pezzi su cento lavorati dalla macchina M_3 sono difettosi)
- I pezzi che arrivano al controllore di qualità sono forniti al 15% dalla macchina M_1 al 35% dalla macchina M_2 e al 50% dalla macchina M_3

Definiamo gli eventi:

1. B_i = “provenienza pezzo” $i = M_1, M_2, M_3$
2. A = “pezzo difettoso”

Teorema della PROBABILITÀ TOTALE

$$P(A) = \text{probabilità “pezzo difettoso”}$$

$$P(A) = P(A|B_{M1})P(B_{M1}) + P(A|B_{M2})P(B_{M2}) + P(A|B_{M3})P(B_{M3})$$

$$P(A) = P(0.01)P(0.15) + P(0.05)P(0.35) + P(0.02)P(0.5) = 0.029$$

REGOLA BAYES:

$P(B_{M2}|A) =$ probabilità che il pezzo difettoso, scelto casualmente tra i pezzi difettosi, proviene dalla M_2

$$P(B_{M2} | A) = \frac{P(A | B_{M2})P(B_{M2})}{P(A)} = \frac{0.05 * 0.35}{0.029} \cong 0.51$$

“Conditional Independence” e grafi (di conoscenza)

DEFINIZIONE:

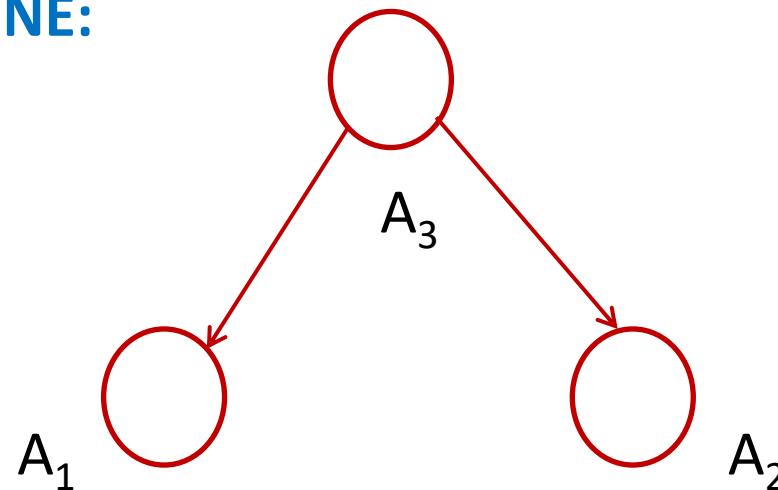
$$P(A_1 \cap A_2 | A_3) \doteq P(A_1 | A_3) \cdot P(A_2 | A_3)$$

DI CONSEGUENZA:

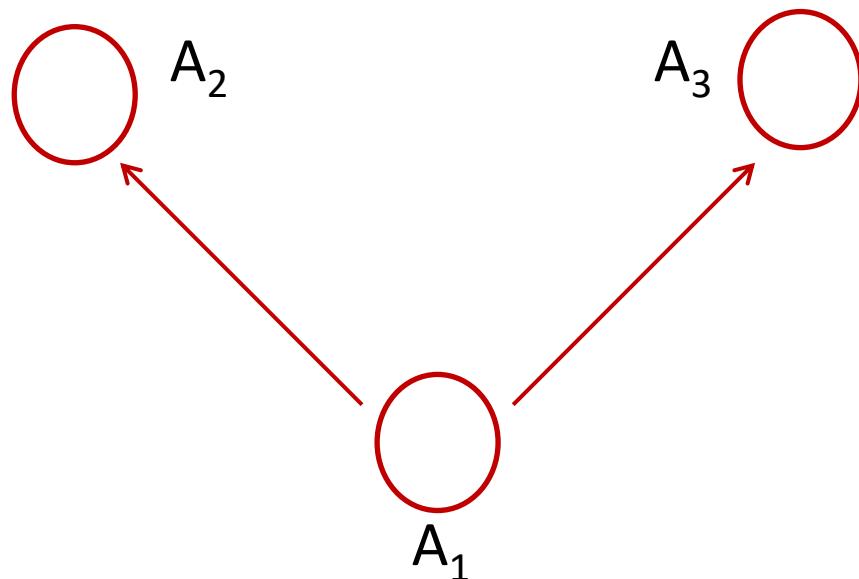
vedi probabilità congiunta e dipendenza

$$\begin{aligned} P(A_1 \cap A_2 \cap A_3) &= P(A_1 \cap A_2 | A_3) \cdot P(A_3) \\ &= P(A_1 | A_3) \cdot P(A_2 | A_3) \cdot P(A_3) \end{aligned}$$

RAPPRESENTAZIONE:



“Conditional Independence” e grafi



OSSERVAZIONE:

NO arco tra
A2 e A3

$$\begin{aligned} P(A_1 \cap A_2 \cap A_3) &= P(A_1 | A_2 \cap A_3) \cdot P(A_2 \cap A_3) \\ &= P(A_1 | A_2 \cap A_3) \cdot P(A_2) \cdot P(A_3) \end{aligned}$$

ALLORA:

$$\begin{aligned} P(A_2 \cap A_3 | A_1) &= P(A_2 \cap A_3) / P(A_1) \\ &= P(A_2) \cdot P(A_3) / P(A_1) \neq P(A_2) \cdot P(A_3) \end{aligned}$$

A2 and A3 are INDEPENDENTS but are NOT CONDITIONAL INDEP

ESEMPIO

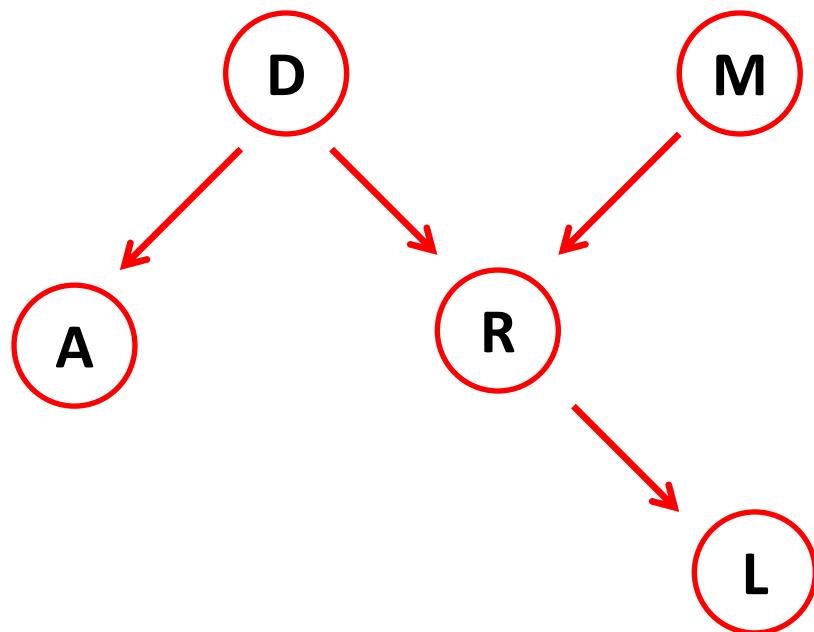
L := “La Lezione comincia alle 10 e 50”

R := “Il docente arriva in Ritardo”

A := “Argomento della lezione è MODELLI su GRAFI.”

M := “Forte Maltempo in atto”

D := “Il Docente è Legato”



ASSUNZIONI:

D INDIP da M

A è DIP da D

R è DIP da M

L è DIP da R

Affidabilità come probabilità di successo

Sulla base dei primi concetti di analisi probabilistica, l'affidabilità (*Reliability*) di un sistema può essere definita come la probabilità che esso risulti funzionante (successo) all'osservatore casuale che "ispeziona" il sistema in un istante di tempo che non ha importanza. L'osservatore casuale è in grado di rilevare lo stato di buono o cattivo funzionamento di tutti i componenti, ripetendo l'osservazione dello stato di ognuno di essi in tempi trascurabili.

Esempio: Controllo del funzionamento del server di prenotazioni a 00:00

Si consideri l'evento

$A_i \hat{=} \text{"componente } i\text{-mo funziona"}$, $i=1..n$,

Affidabilità del sistema seriale

$$R_S \hat{=} P(\text{"sistema funziona"}) \\ = P(\text{"tutti i componenti funzionano"})$$

allora

$$R_S = P(A_1 \cap A_2 \cap \dots \cap A_n) = \prod_{i=1}^n R_i, \quad \text{ove } R_i \hat{=} P(A_i)$$

se aggiungiamo l'ipotesi che i componenti sono indipendenti.

appunto congiunta
che dei singoli
componenti funzionanti.
Se un componente non
funziona, non funziona
tutto il sistema.

Affidabilità del sistema parallelo:

$$R_P = 1 - P(\text{"il sistema non funziona"}) \\ = 1 - P(\text{"tutti i componenti non funzionano"}) \\ = 1 - P(\overline{A}_1 \cap \overline{A}_2 \cap \dots \cap \overline{A}_n) \\ = 1 - \prod_{i=1}^n P(\overline{A}_i) = 1 - \prod_{i=1}^n (1 - P(A_i)) = 1 - \prod_{i=1}^n (1 - R_i)$$

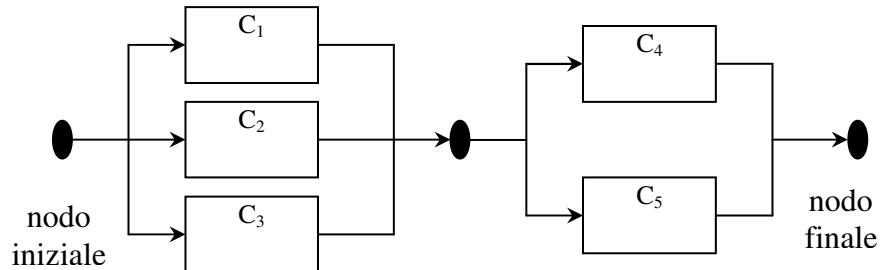
Il sistema smette di funzionare
se non vi è alcun, componente
funzionante.

se aggiungiamo l'ipotesi che i componenti sono indipendenti.

In entrambi i casi si hanno componenti indipendenti!

Diagrammi a Blocchi in Affidabilità

Schematizzazione a blocchi dell'organizzazione di un sistema:



PER DEFINIZIONE:

Il sistema funziona se la numerosità e la disposizione dei componenti rilevati funzionanti dall'osservatore casuale sono tali da garantire almeno un cammino orientato (senza cicli interni) tra nodo iniziale e nodo finale sul quale tutti i componenti sono funzionanti.

Ad esempio: C1, C2 e C5 funzionanti \Rightarrow sistema funziona (due cammini)

Ragionando in termini di cammini, l'affidabilità di un sistema si esprime come segue:

$$P(\text{"Sistema funziona"}) = P\left(\bigcup_{i=1}^n \text{"Tutti funz. sul Camm}_i\text{"}\right)$$

Il fatto che lo stesso componente possa appartenere a più di un cammino esclude la possibilità di trattare gli eventi "Ai" = "Tutti funzionanti sul cammino i", $i=1,\dots,n$ come eventi disgiunti. Quindi per calcolare l'affidabilità del sistema, occorre applicare la formula della probabilità dell'unione di n eventi non disgiunti:

$$\begin{aligned} P\left(\bigcup_{i=1}^n A_i\right) &= P(A_1 \cup A_2 \cup \dots \cup A_n) = \\ &= \sum_{\forall i} P(A_i) - \sum_{1 \leq i < j \leq n} P(A_i \cap A_j) + \sum_{1 \leq i < j < k \leq n} P(A_i \cap A_j \cap A_k) + \dots + (-1)^{n-1} P(A_1 \cap A_2 \cap \dots \cap A_n) \end{aligned}$$

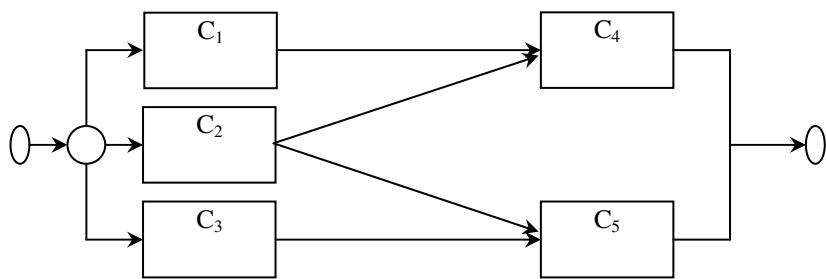
vista a pagina 16

Il metodo dello spazio degli eventi

Quando un sistema è organizzato in una forma che non può essere ricondotta alla composizione di blocchi seriali o paralleli o misti si può fare ricorso al metodo di analisi dello spazio degli eventi. Dove occorre:

- *Individuare tutti i possibili eventi;*
- *Definire gli eventi corrispondenti al funzionamento del sistema;*
- *Sviluppare le formule di probabilità.*

Analisi del sistema:



Considerati i soliti eventi

$$A \triangleq \text{"sistema funziona"}$$

$$A_i \triangleq \text{"componente } i\text{-mo funziona"}$$

l'analisi è basata sulla relazione che lega l'evento A agli eventi cui corrisponde il funzionamento del sistema e che vanno individuati per ispezione del diagramma a blocchi. Per il sistema considerato, risulta:

$$A = (A_1 \cap A_4) \cup (A_2 \cap A_4) \cup (A_2 \cap A_5) \cup (A_3 \cap A_5)$$

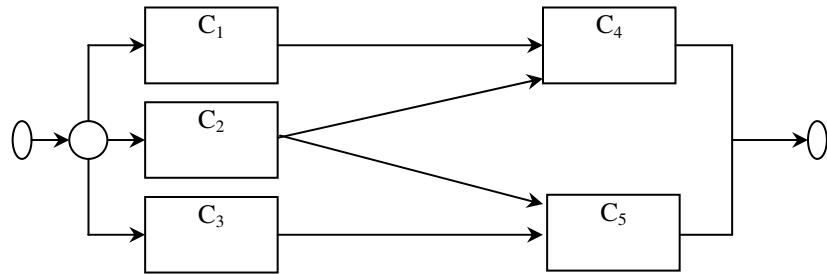
Osservando che i termini dell'unione hanno intersezione non nulla a due a due (A_4 compare nel primo e nel secondo termine, ad esempio), l'affidabilità del sistema può essere calcolata con la formula seguente:

$$\begin{aligned} P(A) &= P(E_1 \cup E_2 \cup E_3 \cup E_4) = \\ &= \sum_{1 \leq i \leq 4} P(E_i) - \sum_{1 \leq i < j \leq 4} P(E_i \cap E_j) + \sum_{1 \leq i < j < k \leq n} P(E_i \cap E_j \cap E_k) - P(E_1 \cap E_2 \cap E_3 \cap E_4) \end{aligned}$$

dove $E_1 \triangleq A_1 \cap A_4$, $E_2 \triangleq A_2 \cap A_4$ e così via.

Esempio: Probabilità totale e analisi di affidabilità

Alternativa al metodo dello spazio degli eventi per il sistema:



Considerati i due eventi:

$B \hat{=} \text{"Componente 2 funziona"}$ e $A \hat{=} \text{"Sistema funziona"}$

una possibile partizione di Ω è $\{B, \bar{B}\}$ e quindi la probabilità che il sistema sia funzionante è:

$$\begin{aligned} &\text{dalla formula della totale} \\ P(A) &= P(A \cap B) + P(A \cap \bar{B}) \\ &= P(A | B)P(B) + P(A | \bar{B})P(\bar{B}) \end{aligned}$$

dopo aver calcolato: $P(A | B)$ e $P(A | \bar{B})$.

Con la condizione di funzionamento certo del componente 2, il sistema si riduce al parallelo dei componenti 4 e 5. Applicando appunto l'affidabilità in parallelo

$$\text{Allora: } P(A | B) = 1 - (1 - R_4)(1 - R_5)$$

Viceversa, assumendo che il componente 2 non funzioni, il sistema si riduce al parallelo di due blocchi; il primo blocco essendo costituito dalla serie dei componenti 1 e 4 e il secondo blocco dalla serie dei componenti 3 e 5. Due blocchi indipendenti tra loro!

$$\text{Allora: } P(A | \bar{B}) = 1 - (1 - R_1R_4)(1 - R_3R_5)$$

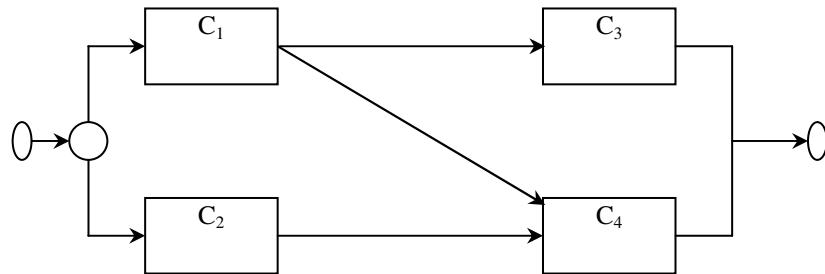
In definitiva:

$$R_{\text{SISTEMA}} = [1 - (1 - R_4)(1 - R_5)]R_2 + [1 - (1 - R_1R_4)(1 - R_3R_5)](1 - R_2)$$

CONGIUNTA

Esempio: Spazio degli eventi, probabilità totale e regola di Bayes

È dato il seguente schema a blocchi dell'organizzazione di un sistema:



ed è nota l'affidabilità dei singoli componenti, intesa come “probabilità di successo”:

$$R_{C1} = 0,95, R_{C2} = 0,85, R_{C3} = 0,92, R_{C4} = 0,98$$

- Calcolare l'affidabilità del sistema applicando il metodo dello spazio degli eventi;
- Ripetere il calcolo dell'affidabilità facendo uso della formula della probabilità totale;
- Calcolare la probabilità che il componente C1 funzioni, posto che il sistema funziona.

Soluzione

Si definisca l'evento $A \hat{=} \text{almeno un cammino funzionante nel sistema}$.

I cammini possibili sono:

$$B_1 = C_1 \cap C_3, B_2 = C_2 \cap C_4, B_3 = C_1 \cap C_4 .$$

Quindi,

$$A = B_1 \cup B_2 \cup B_3 = (C_1 \cap C_3) \cup (C_2 \cap C_4) \cup (C_1 \cap C_4)$$

$$\begin{aligned} P(A) &= P(B_1 \cup B_2 \cup B_3) \\ &= P(B_1) + P(B_2) + P(B_3) - P(B_1 \cap B_2) - P(B_1 \cap B_3) - P(B_2 \cap B_3) + \\ &\quad + P(B_1 \cap B_2 \cap B_3) \\ &= P(B_1) + P(B_2) + P(B_3) - P(B_1 \cap B_2) - P(B_1 \cap B_3) + \\ &\quad + P(B_1 | B_2 \cap B_3) * P(B_2 \cap B_3) \\ &= P(B_1) + P(B_2) + P(B_3) - P(B_1 \cap B_2) - P(B_1 \cap B_3) + \\ &\quad + P(B_1 | B_2 \cap B_3) * P(B_2 | B_3) * P(B_3) \end{aligned}$$

$$\begin{aligned}
&= (R_{C1} * R_{C3}) + (R_{C2} * R_{C4}) + (R_{C1} * R_{C4}) - (R_{C1} * R_{C3} * R_{C2} * R_{C4}) + \\
&\quad - (R_{C2} * R_{C1} * R_{C4}) - (R_{C3} * R_{C1} * R_{C4}) + (R_{C3} * R_{C2} * R_{C1} * R_{C4}) \\
&= 0,874 + 0,833 + 0,931 - 0,728042 - 0,79135 - 0,85652 + 0,728042 \\
&= 0,99013
\end{aligned}$$

Usando la formula della probabilità totale, si considerino i seguenti eventi:

$A \hat{=} \text{il sistema funziona}$

$B \hat{=} \text{il componente } C_1 \text{ funziona}$

allora

$$P(A) = P(A | B) * P(B) + P(A | \bar{B}) * P(\bar{B}) = P(A \cap B) + P(A \cap \bar{B}).$$

Se il componente C_1 non funziona, il sistema si riduce alla serie fra C_2 e C_4

$$P(A | \bar{B}) = R_{C2} * R_{C4}.$$

Se il componente C_1 funziona, il sistema si riduce al parallelo fra i componenti C_3 e C_4

$$P(A | B) = 1 - [(1 - R_{C3}) * (1 - R_{C4})].$$

In conclusione

$$P(A) = [1 - (1 - R_{C3}) * (1 - R_{C4})] * R_{C1} + (R_{C2} * R_{C4}) * (1 - R_{C1}) = 0,99013$$

La probabilità che il componente C_1 funzioni, posto che il sistema funzioni, si ottiene applicando la regola di Bayes:

$A \hat{=} \text{il sistema funziona}$

$B \hat{=} \text{il componente } C_1 \text{ funziona}$

$$P(B | A) = \frac{P(A | B) * P(B)}{P(A)}$$

$$P(A | B) * P(B) = [1 - (1 - R_{C3}) * (1 - R_{C4})] * R_{C1}$$

Quindi

$$P(B | A) = \frac{0,94848}{0,99013} = 0,957934816 \approx 0,96$$

Esempio: Probabilità totale e limite superiore all'Affidabilità

Sia $\{C_1, C_2, \dots, C_n\}$ l'insieme dei cammini individuati su un diagramma a blocchi assegnato e si consideri l'evento:

$$D_i \hat{=} \text{"Almeno un componente in } C_i \text{ è guasto".}$$

In linea di principio si può scrivere:

$$\begin{aligned} 1 - R &= P(D_1 \cap D_2 \cap \dots \cap D_n) \\ &= P(D_n | D_1 \cap D_2 \cap \dots \cap D_{n-1}) \dots P(D_2 | D_1)P(D_1) \end{aligned}$$

ma nella pratica, avendo a che fare con diagrammi di dimensioni non banali, la relazione precedente non è utilizzabile.

Però si potrebbe calcolare un valore limite (superiore), rinunciando al valore vero.

Per ricavare il valore limite vero occorre provare che:

$$P(D_2 | D_1) \geq P(D_2)$$

PROVA:

$$P(D_2) = P(D_2 | D_1)P(D_1) + P(D_2 | \overline{D_1})(1 - P(D_1))$$

e riconoscendo che

$$\begin{aligned} P(D_2 | \overline{D_1}) &= \text{Prob "almeno un componente guasto in } C_2 | \text{ tutti funzionano in } C_1 \\ &\text{risulta} \end{aligned}$$

$$\begin{aligned} P(D_2 | \overline{D_1}) &= 1 - \prod_{\substack{i \in C_2 \\ i \notin C_1}} R_i \leq 1 - \prod_{i \in C_2} R_i = P(\overline{D_2}) \\ \Rightarrow P(D_2) &\leq P(D_2 | D_1)P(D_1) + P(D_2)(1 - P(D_1)) \\ \Rightarrow 0 &\leq P(D_2 | D_1) - P(D_2) \end{aligned}$$

Quindi

$$P(D_i | D_1 \dots D_{i-1}) \geq P(D_i) \Rightarrow (1 - R) \geq \prod_i P(D_i)$$

In conclusione:

$$R \leq 1 - \prod_{i=1}^n \left(1 - \prod_{k \in C_i} R_k \right)$$

IMPORTANTE

Le prove di Bernoulli

Rappresentano un possibile modello della ripetizione di "n" tentativi di accesso di un utente ad una risorsa, permettendo di calcolare: ad esempio richiesta di usare cpu, se fallisco ritento

- La probabilità di "k" successi su "n" tentativi;
- La probabilità che il primo successo si verifichi al "k-mo" tentativo.

Lo spazio dei possibili risultati per una sola prova è: $\Omega_1 = \{0,1\}$;

Per due prove è: $\Omega_2 = \{(0,0), (0,1), (1,0), (1,1)\}$

Per "n" prove è: $\Omega_n = \{2^n \text{ n-più di zero e uno}\}$

Nel modello di Bernoulli, l'assegnamento delle probabilità ai risultati di Ω_1 è:

$$P(0) = q \geq 0, \quad P(1) = p \geq 0 \quad \text{con} \quad p + q = 1$$

perciò successi p, fallimenti q=1-p

Per l'assegnamento delle probabilità a Ω_n è necessario definire:

$A_i \hat{=} \text{"successo alla prova i-ma"}, \quad \overline{A_i} \hat{=} \text{"insuccesso alla prova i-ma"}$.

Si consideri ora un possibile $\omega_n \in \Omega_n$: sequenza di prove

$$\omega_n = A_1 \cap A_2 \cap \dots \cap A_k \cap \overline{A_{k+1}} \cap \dots \cap \overline{A_n}.$$

Nel modello di Bernoulli, le prove sono indipendenti e le probabilità di successo e insuccesso non cambiano all'aumentare delle prove, e l'ordine dei successi/insuccessi è irrilevante. Le prove sono n e devono essere tutte eseguite, quando si vuole che la probabilità di "k" successi su "n" tentativi. Dunque si ottiene:

$$P(\omega_n) = P(A_1 \cap A_2 \cap \dots \cap A_k \cap \overline{A_{k+1}} \cap \dots \cap \overline{A_n}) = p^k q^{n-k}.$$

Come caso particolare si osserva che:

$$q^{k-1} p$$

è la probabilità che il primo successo si realizzi al tentativo "k-simo" (nota anche come probabilità geometrica). Da notare che le prove a disposizione diventano illimitate, quando si vuole la probabilità che il primo successo si verifichi al "k-mo" tentativo

Inoltre, se non interessa l'ordine in cui si ottengono:

$$P("k \text{ successi in } n \text{ prove}") = \binom{n}{k} p^k q^{n-k} \quad k = 0, 1, \dots, n .$$

La precedente è nota come **probabilità binomiale**.

Non interessa l'ordine dei successi

Generalizzazione delle prove di Bernoulli

Consiste nell'associare "k" possibili risultati ad ognuna delle prove.

Esempio: a "k" possibili risorse può essere mirato ognuno degli "n" tentativi di accesso di un utente, oppure l'unico tentativo di accesso di "n" utenti che si susseguono uno dopo l'altro.

Occorre identificare lo spazio dei risultati di n prove generalizzate (Ω_n) .

Sia ω l'evento elementare di Ω_n , allora:

$$\omega \triangleq \left\{ \begin{array}{c} 1,1,\dots,1; 2,2,\dots,2; \dots; k,k,\dots,k \\ n_1 \text{ volte} \quad n_2 \text{ volte} \quad \dots \quad n_k \text{ volte} \end{array} \right\} \text{ e } (n_1 + n_2 + \dots + n_k = n)$$

$$P(n_1, n_2, \dots, n_k) = \binom{n}{n_1} \binom{n-n_1}{n_2} \dots \binom{n-n_1-\dots-n_{k-1}}{n_k} P_1^{n_1} P_2^{n_2} \dots P_k^{n_k}$$

dove

$$Prob\{ \text{"vero il j-mo risultato nella generica prova"} \} \triangleq P_j$$

Dunque in definitiva nelle prove di Bernoulli la probabilità di successo rimane costante e indipendente!

La probabilità Binomiale Negativa

Se si ribaltano i termini dell'esperimento bernoulliano, fissando il numero di successi (k) che si vogliono ottenere e lasciando libero il numero di tentativi (n), si vuole costruire un modello di ripartizione per determinare la probabilità di quanti tentativi bisogna effettuare per ottenere k successi. Si definiscano pertanto i seguenti eventi:

- $A = \text{"n° tentativi per realizzare k successi"}$. Con k fissato a priori e n° tentativi (n) libero, purché sia $n \geq k$
- $B = \text{"esattamente } k-1 \text{ successi in } n-1 \text{ tentativi"}$ successi fissati quindi prove di Bernoulli!
- $C = \text{"il tentativo } n\text{-esimo è un successo"}$

L'evento A può essere calcolato a partire dai due eventi indipendenti B e C. A è verificato se si verificano gli eventi B e C, ossia:

$$A = B \cap C$$

Quindi possiamo calcolare la probabilità che si verifichi l'evento A come:

$$P(A) = P(B \cap C) = P(B | C) * P(C)$$

Poiché gli eventi B e C sono tra loro indipendenti, possiamo riscrivere la predente espressione come segue:

$$P(A) = P(B) * P(C)$$

P(B) si può calcolare con il modello di Bernoulli. Abbiamo infatti visto che, nel caso in cui non sia importante la sequenza dei successi, la probabilità di ottenere esattamente k successi in n prove è data dalla probabilità binomiale:

$$P(\text{"k successi in n prove"}) = \binom{n}{k} p^k q^{n-k} \quad k = 0, 1, \dots, n.$$

P(B) si calcola considerando: una particolare sequenza di n-1 tentativi, in cui i primi k-1 sono successi ed i restanti $n-1-(k-1)=n-k$ sono insuccessi, e moltiplicando la probabilità di quella sequenza per il numero di tutte le possibili sequenze di n-1 tentativi con k-1 successi.

$$P(B) = P(B) \cdot P(C) = \binom{n-1}{k-1} p^{k-1} q^{n-k}$$

La probabilità di successo di ogni singola prova è nota ed è pari a p . La probabilità che si verifichi l'evento C è quindi: $P(C) = p$. A questo punto è possibile calcolare la probabilità dell'evento A per semplice sostituzione delle probabilità calcolate.

$$P(A) = \binom{n-1}{k-1} p^{k-1} q^{n-k}$$

La precedente è nota come *probabilità binomiale negativa* (o di Pascal).

Si può notare come diremo molto in futuro che la binomiale è la somma di prove di Bernoulli.

Prove di Bernoulli in Affidabilità: Sistemi “m-out-of-n”

Per questo tipo di sistemi l'affidabilità è definita come segue

$$R_{m|n} = P(\text{"almeno } m \text{ componenti funzionano"}) = P(A(\geq m))$$

Partizionando lo spazio dei risultati Ω si può scrivere

Cioè il sistema funziona quando almeno “m” componenti su “n” funzionano. Cioè evento A($\geq m$)

$$\begin{aligned} R_{m|n} &= P\left(\bigcup_{i=m}^n \text{"esattamente } i \text{ componenti funzionano"}\right) \\ &= \sum_{i=m}^n P(\text{"esattamente } i \text{ componenti funzionano"}) \end{aligned}$$

Dunque: $R_{m|n} = P(A(\geq m)) = P\left(\bigcup_{i=m}^n A(=i)\right) = \sum_{i=m}^n P(A(=i))$ Rifacendosi alle prove di Bernoulli:

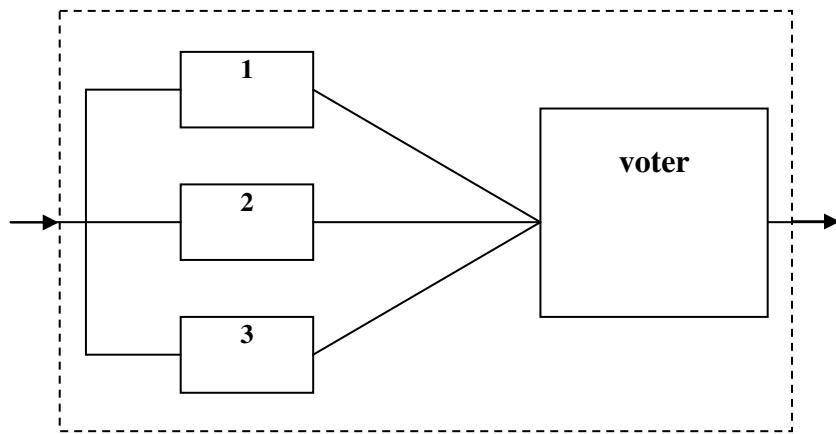
$$R_{m|n} = \sum_{i=m}^n \binom{n}{i} R^i (1-R)^{n-i}$$

con ipotesi che i componenti sono indipendenti ed identici con $AFF=R$.

Ridondanza con voto di maggioranza

Come esempio di sistema “m-out-of-n” si riporta il sistema a ridondanza modulare tripla (TMR). L’organizzazione TMR è classica nei sistemi di elaborazione e di controllo di apparecchiature o impianti. Tre oggetti, solitamente identici, lavorano in parallelo producendo un risultato che viene mandato in ingresso ad un comparatore (voter). Il comparatore ha il compito di rilevare l’eventuale disaccordo fra i risultati, che in condizioni di normalità devono essere identici assicurando che il risultato è corretto. Scoprire che uno degli oggetti produce un risultato diverso dagli altri induce a ritenere che sia proprio quello a funzionare male: in tal caso quell’oggetto viene messo fuori linea. Finché i due oggetti rimanenti continueranno a produrre lo stesso risultato si potrà continuare a ritenere che sia quello corretto. Al primo disaccordo il sistema dovrà essere considerato mal funzionante e quindi disattivato.

Se non sono identici il comparatore non funziona!



L'affidabilità del TMR è:

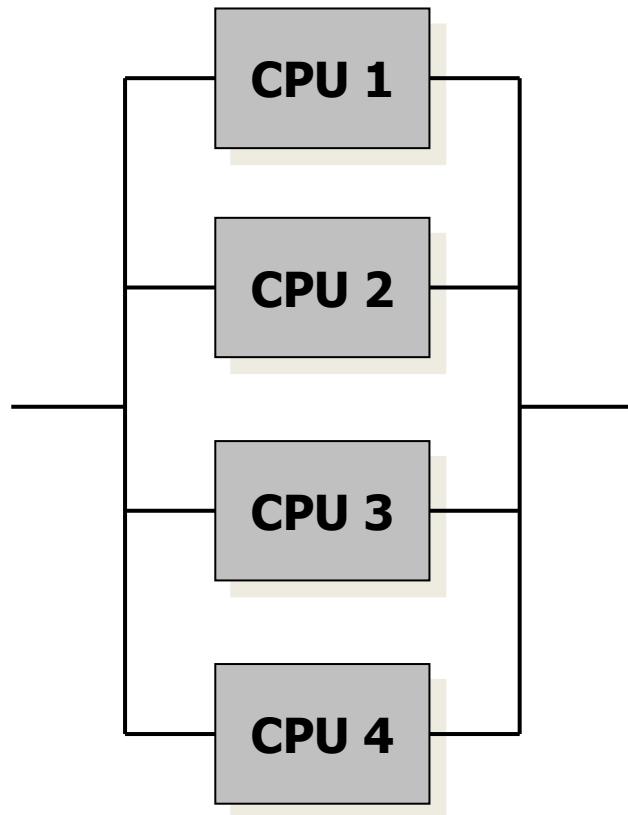
$$R_{m|n} = \sum_{i=m}^n \binom{n}{i} R^i (1-R)^{n-i}$$

dunque sostituendo n=3 ed m=2

$$R_{TMR} = \binom{3}{2} R^2 (1-R)^{3-2} + \binom{3}{3} R^3 (1-R)^{3-3}$$

Esercizio su sistema “m-out-of-n”: descrizione del problema e quesiti (1|2)

- Un sistema di calcolo ad elaborazione parallela, dotato di 4 CPU identiche e una RAM comune, è considerato "funzionante" se risultano funzionanti almeno 2 qualsivoglia delle 4 CPU, a parte la RAM.
- Sapendo che l'affidabilità della singola CPU vale 0.9, proporre un modello di calcolo e calcolare l'affidabilità del sistema.



Esercizio: descrizione del problema e quesiti (2|2)

1. Qual è la probabilità che il sistema lavori al massimo della sua prestazione, ovvero con tutte e quattro le CPU funzionanti?
2. Qual è la probabilità che "funzioni al minimo, cioè con 2 sole CPU", però dando per certo che funziona?
3. Ricalcolare l'affidabilità del sistema nell'ipotesi che debba funzionare una specifica CPU (ad es. la n°1) e poi anche solo un'altra delle rimanenti 3.
4. Ricalcolare l'affidabilità del sistema dando per certo che funziona una specifica CPU (ad es. la la n°1).
5. Rappresentare il diagramma a blocchi del sistema e provare a calcolare l'affidabilità col metodo dello spazio degli eventi.

Esercizio: modello e calcolo dell'affidabilità

Si tratta di un sistema «m-out-of-n» con n=4 e m=2.

Ricordando:

$$R_{m|n} = \sum_{i=m}^n \binom{n}{i} \cdot (R)^i \cdot (1-R)^{n-i}$$

Si calcola:

$$R_{2|4} = \sum_{i=2}^4 \binom{4}{i} \cdot (0.9)^i \cdot (0.1)^{4-i} = 0.0486 + 0.2916 + 0.6561 = 0,9963$$

Osservazione:

Sia C_i l'evento: «esattamente i cpu funzionano»

e sia F l'evento: «il sistema funziona»

Allora:

$$R_{2|4} \doteq P(F) = P\left(\bigcup_{i=2}^{i=4} (F \cap C_i)\right) = \sum_{i=2}^{i=4} P(F \cap C_i)$$

Perciò:

$$P(F \cap C_2) = 0.0486; \quad P(F \cap C_3) = 0.2916; \quad P(F \cap C_4) = 0.6561$$

Esercizio: risposte ai quesiti 1. e 2.

Quesito 1:

$$P(F \cap C_4) = \binom{4}{4} (0.9)^4 (0.1)^0 = 0.6561$$

Quesito 2:

Sia E l'evento: «solo 2 su 4 funzionano | sistema funziona»

Allora:

$$P(E) = \frac{P(F \cap C_2)}{P(F)} = \frac{0.0486}{0.9963} = 0.0487$$

Osservazione:

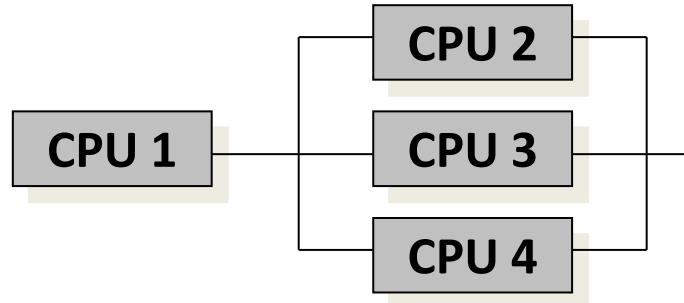
La probabilità (incondizionata) che il sistema funzioni al livello di prestazione 2 su 4 è quasi la stessa della probabilità (condizionata) che il sistema funzioni allo stesso livello quando funziona, solo perché $P(F)$ è quasi =1.

Ma le due probabilità in questione sono concettualmente ben diverse!

Esercizio: risposte ai quesiti 3. e 4.

Quesito 3:

Schema
corrispondente:

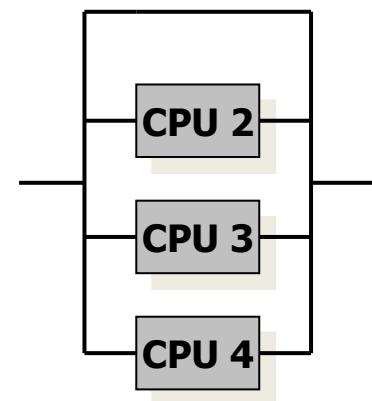


$$\begin{aligned} R &= R_{CPU1} \cdot R_{1|3} = R_{CPU1} \cdot (1 - (1 - R_2)(1 - R_3)(1 - R_4)) \\ &= 0.9 \cdot (1 - 0.9)(1 - 0.9)(1 - 0.9) = 0.8991 \end{aligned}$$

Quesito 4:

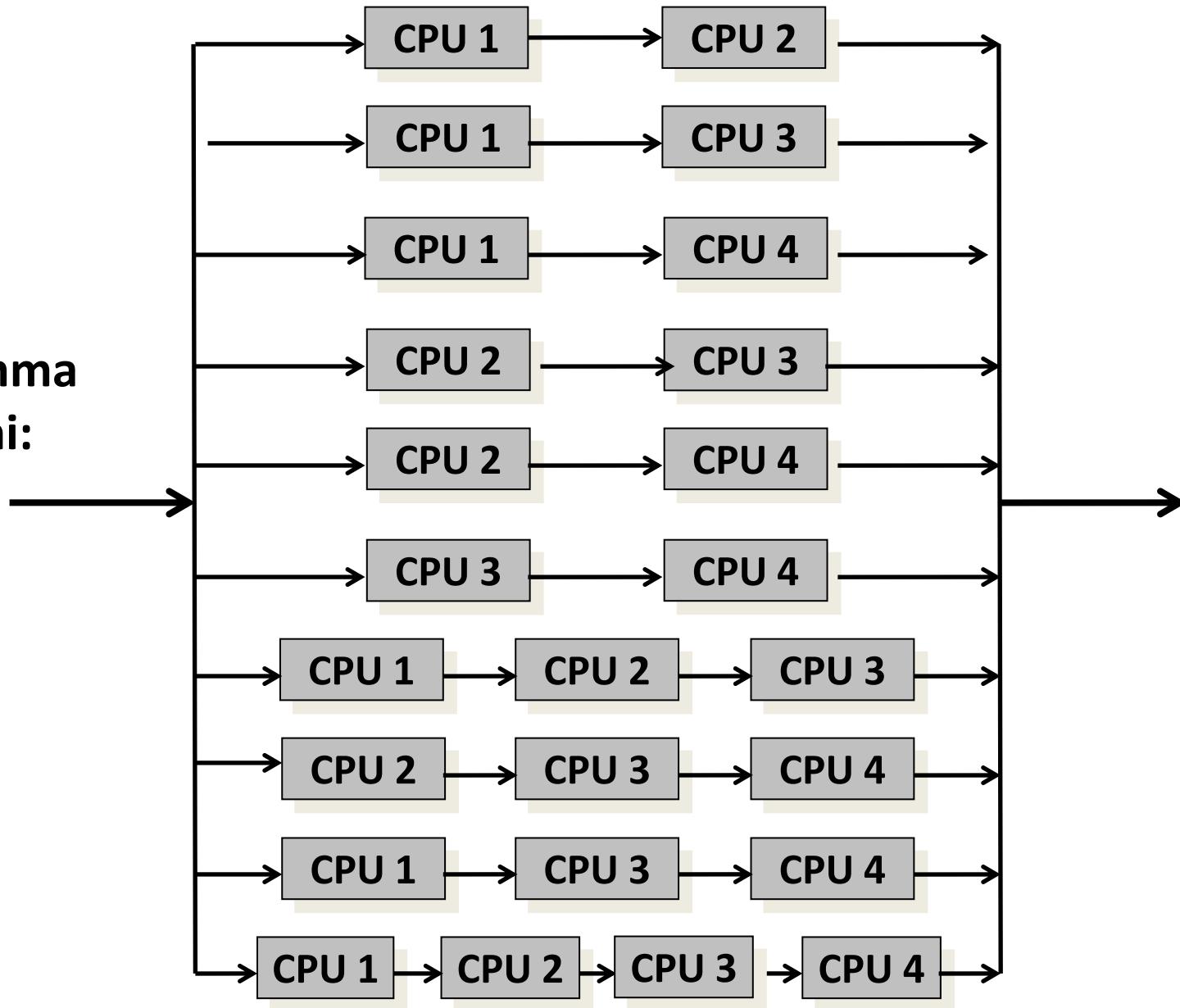
$R = P(2\text{-out-of-4} \mid \text{CPU 1 funziona certamente}) = P(1\text{-out-of-3})$

$$R = R_{2|4} \mid \text{CPU1} = R_{1|3} = 1 - (1 - R_2)(1 - R_3)(1 - R_4) = ???$$



Esercizio: risposta al quesito 5.

Diagramma
a blocchi:



Esercizio 1: descrizione del problema e quesiti (1|2)

Un sistema di elaborazione è composto da 4 processori che possono accedere tramite 1 bus a 4 moduli di memoria.

- Ad ogni ciclo di clock, ognuno dei processori può richiedere il bus per accedere ad uno dei moduli di memoria.
- Se più di uno è richiedente, il conflitto viene risolto assegnando il bus in maniera completamente casuale (ovvero equiprobabile).
- Un processore richiedente, al quale non è stato assegnato il bus, può anche non insistere nella richiesta al successivo ciclo di clock.
OVVERO INDIPENDENZA TRA GLI EVENTI
- La probabilità che un certo numero « n , ($n=0,1,\dots,4$)» di processori richieda di accedere al bus nello stesso ciclo di clock è nota, per ipotesi, dalla tabella seguente:

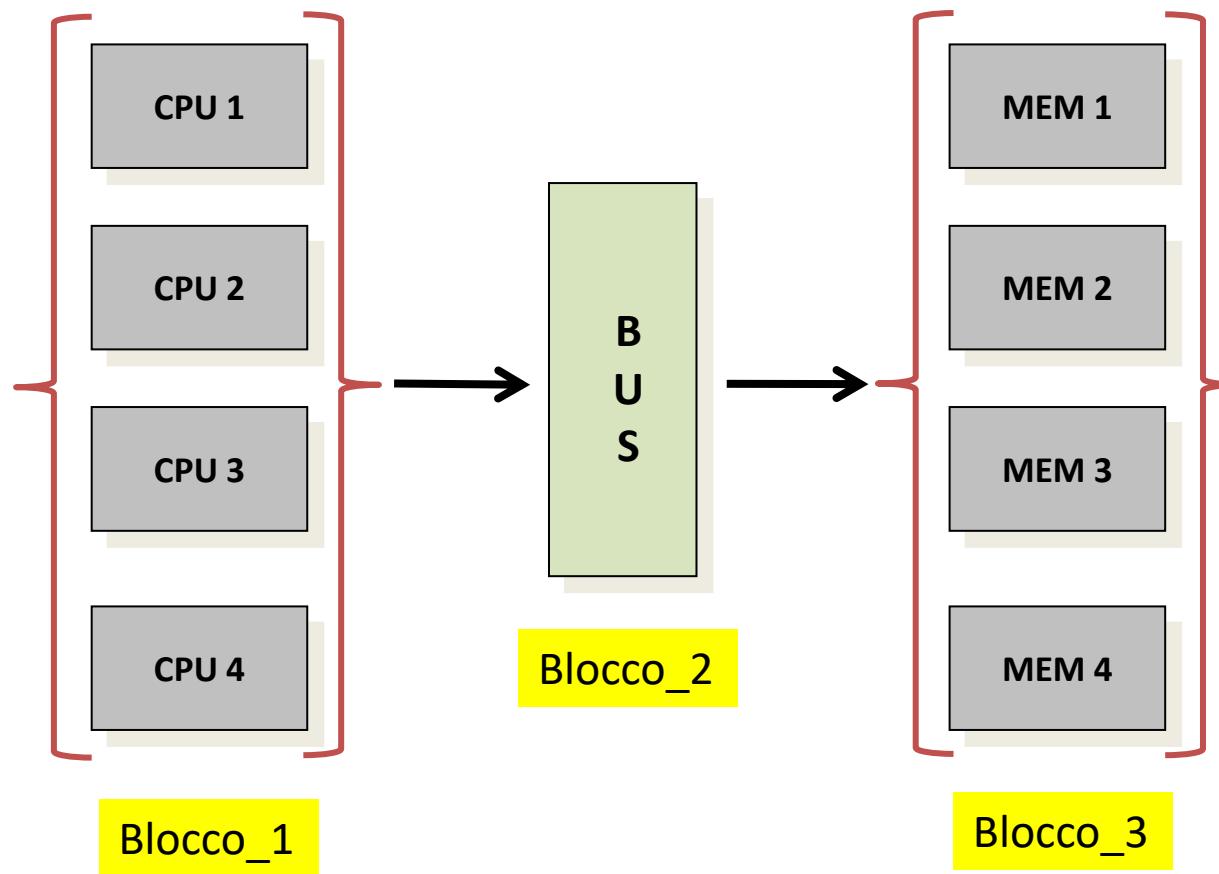
n° proc. richied.	0	1	2	3	4
probabilità	0,2	0,40	0,20	0,12	0,08

Esercizio 1: descrizione del problema e quesiti (2|2)

Indicare il modello di Affidabilità e poi calcolare:

1. la probabilità che si verifichi conflitto per il bus; quindi 2,3 o 4 processori richiedano il bus
2. la **probabilità che uno specifico processore che ne abbia fatto richiesta si veda assegnato il bus;** Bayes ?
3. la probabilità che quel processore ottenga il bus, per la prima volta, alla sua terza richiesta; prove di Bernoulli
4. la probabilità che occorrono più di 3 richieste per ottenere il bus per la prima volta;
5. la **probabilità che lo stesso processore debba fare 4 richieste per ottenere il bus per 2 volte,** sia pure non necessariamente consecutive. 2 successi in 4 tentativi: binomiale negativa

Esercizio 1: modello per il calcolo dell'affidabilità



IPOTESI:

Il sistema funziona quando funzionano almeno 1 CPU, il BUS e almeno 2 MEM

Modello risultante:

$$R_{Sis} = R_{B1} \cdot R_{B2} \cdot R_{B3} = R_{1|4} \cdot R_{BUS} \cdot R_{2|4}$$

Esercizio 1: risposta al quesito 1.

- Sia B_n l'evento: «n ($n=0,1,\dots,4$) processori richiedono il bus nello stesso clock»
- Sia E l'evento d'interesse: «almeno 2 processori nello stesso clock, ovvero conflitto»

Allora, poiché gli eventi B_n sono disgiunti, cioè tra loro incompatibili comunque li si cerchi di raggruppare (a 2 a 2, a 3 a 3 e tutti insieme), risulta:

$$P(E) = P\left(\bigcup_{n=2}^{n=4} B_n\right) = \sum_{n=2}^{n=4} P(B_n) = 0,20 + 0,12 + 0,08 = 0,40$$

Osservazioni:

Dalla tabella fornita con il problema si evince che gli eventi $[B_n, n=0,1,\dots,4]$ sono eventi diciamo elementari, in quanto le relative probabilità sono assegnate!

Gli eventi $[B_n, n=0,1,\dots,4]$ avrebbero potuto non essere elementari, ma calcolabili a partire da altri eventi elementari, cioè direttamente osservabili, ad es. $[C_i, i=1,\dots,4]$, con $C_i = \text{«il processore } i\text{-esimo richiede il bus»}$.

In tal caso, sarebbe sorta una difficoltà: $P(C_i | C_j, C_k, C_l) = ? \dots$

Come si può ridurre una tale difficoltà? Come si può eliminare del tutto?

Esercizio 1: risposta al quesito 2.

Sia A l'evento (d'interesse): «un processore richiedente ottenga il bus»

Ricorrendo al concetto di probabilità condizionata, a partire dalle probabilità note degli eventi B_n , però con $n=1, \dots, 4$, occorre calcolare le seguenti:

$$P(\hat{B}_n) = \frac{P(B_n)}{\sum_{i=1}^{i=4} P(B_i)} \quad n = 1, \dots, 4 \quad \rightarrow \quad \begin{cases} P(\hat{B}_1) = \frac{0.4}{0.8} = 0.50, & P(\hat{B}_2) = \frac{0.2}{0.8} = 0.25, \\ P(\hat{B}_3) = \frac{0.12}{0.8} = 0.15, & P(\hat{B}_4) = \frac{0.08}{0.8} = 0.10 \end{cases}$$

D'altra parte, l'ipotesi di completa casualità nell'assegnazione in caso di conflitto si traduce nelle seguenti:

$$P(A | \hat{B}_2) = 0.5, \quad P(A | \hat{B}_3) = 0.33, \quad P(A | \hat{B}_4) = 0.25 \quad [\text{a parte: } P(A | \hat{B}_1) = 1.0]$$

A questo punto, si può applicare la formula della probabilità totale:

$$\begin{aligned} P(A) &= \sum_{n=1}^{n=4} P(A | \hat{B}_n) \cdot P(\hat{B}_n) = \\ &= 1 \cdot 0.50 + 0.5 \cdot 0.25 + 0.33 \cdot 0.15 + 0.25 \cdot 0.10 = 0.70 \end{aligned}$$

funziona anche Bayes per calcolare la probabilità

Esercizio 1: risposta ai quesiti 3. e 4.

Sia E_3 l'evento (d'interesse): «il bus è ottenuto dal processore richiedente al terzo tentativo»;

Osservazione:

- $P(E_3)$ può essere calcolata come probabilità geometrica sotto le seguenti ipotesi: I tentativi sono indipendenti e ripetibili quanto si vuole; inoltre, la probabilità di successo al generico tentativo è sempre la stessa e pari alla $P(A)=0.7$ calcolata in risposta al quesito 2.

$$\Rightarrow P(E_3) = (1 - P(A)) \cdot (1 - P(A)) \cdot P(A) = 0.3 \cdot 0.3 \cdot 0.7 = 0.063$$

Sia F l'evento (d'interesse): «occorrono più di tre richieste per ottenere il bus per la prima volta»;

- Allora:

$$\begin{aligned} P(F) &= 1 - [P(E_1) + P(E_2) + P(E_3)] \\ &= 1 - [0.7 + 0.21 + 0.063] = 1 - 0.973 = 0.027 \end{aligned}$$

(potendo ripetere la richiesta all'infinito)

Osservazione: E_2 ed E_3 sono eventi tra di loro incompatibili (e pure con E_1) e, comunque, ognuno dei due risulta come «and» di altri eventi indipendenti la cui probabilità si calcola a sua volta come probabilità totale (di eventi dipendenti)!

Esercizio 1: risposta al quesito 5.

Sia $G_{2|4}$ l'evento (d'interesse): «fare 4 richieste per ottenere il bus per 2 volte, sia pure non necessariamente consecutive»;

Osservazione:

- $P(G_{2|4})$ può essere calcolata come una nuova probabilità, detta binomiale negativa, sotto le ipotesi delle prove di Bernoulli, nel caso speciale di potere ripetere le prove all'infinito e per la fortuna che non è stato chiesto un ordine preciso nell'ottenere i 2 successi sulle 4 richieste!

Siano: $G_{k-1|n-1}$ = « $k-1$ successi su $n-1$ prove» e S = «successo alla prova n -sima»

Allora: $P(G_{k|n}) = P(G_{k-1|n-1} | S) \cdot P(S)$, con $P(S) = P(A)$

e con: $P(G_{k-1|n-1} | S) = P(G_{k-1|n-1}) = \binom{n-1}{k-1} P(A)^{k-1} (1 - P(A))^{n-k}$

dunque: $P(G_{n|k}) = \binom{n-1}{k-1} P(A)^k (1 - P(A))^{n-k}$

altro non è che la dimostrazione della binomiale negativa

$$\Rightarrow P(G_{4|2}) = \binom{4-1}{2-1} 0.7^2 (1 - 0.7)^{4-2} = 3 \cdot 0.49 \cdot 0.09 = 0.132$$

Esercizio 2: descrizione del problema

Un'applicazione gestisce l'archivio dei prodotti presenti in un ipermercato e aggiorna i prezzi quotidianamente.

I dati utilizzati dall'applicazione risiedono in 4 diversi database, ognuno dei quali contenente specifiche categorie di prodotti presenti nell'ipermercato.

I database, però, possono essere soggetti a inconsistenza, cioè i prezzi dei vari prodotti non sono sempre aggiornati in modo da riflettere le ultime variazioni occorse.

Database	% di prodotti il cui prezzo non è aggiornato
Db Alimentari	2%
Db Giocattoli	7%
Db Faidate	5%
Db Abbigliam	13%

Esercizio 2: dati, quesiti e risposte (1|2)

La totalità dei prodotti presenti può essere partizionata come segue:

69% alimentari, 6% giocattoli, 2% fai da te e, infine, il 23% abbigliamento.

cosa accade se tutte le percentuali sono al 25%?

Con quale probabilità un “cliente casuale” troverà un prodotto (quasivoglia) il cui prezzo non è aggiornato?

Sia “A” l’evento (d’interesse) : “prezzo prodotto non aggiornato”
E sia “ $A|Db_i$ ”, $i=1,\dots,4$ l’evento condizionato: “prezzo non aggiornato, posto che provenga dal database i-esimo”.

Trattando le percentuali come probabilità, si può calcolare $P(A)$ ricorrendo alla formula della probabilità totale:

$$\begin{aligned} P(A) &= P(A/Db_1)*P(Db_1) + P(A/Db_2)*P(Db_2) + \\ &\quad + P(A/Db_3)*P(Db_3) + P(A/Db_4)*P(Db_4) = 0.049 \end{aligned}$$

Esercizio 2: dati, quesiti e risposte (2|2)

Calcolare la probabilità dei seguenti eventi “Dbi, i=1,...,4” = “prezzo non aggiornato estratto è proveniente dal database i”

$$P(Db1 | A) = \frac{P(A | Db1)P(Db1)}{P(A)} = \frac{0.02 * 0.69}{0.049} = 0.28$$

$$P(Db2 | A) = \frac{P(A | Db2)P(Db2)}{P(A)} = \frac{0.07 * 0.06}{0.049} = 0.086$$

$$P(Db3 | A) = \frac{P(A | Db3)P(Db3)}{P(A)} = \frac{0.05 * 0.02}{0.049} = 0.020$$

$$P(Db4 | A) = \frac{P(A | Db4)P(Db4)}{P(A)} = \frac{0.13 * 0.23}{0.049} = 0.61$$

Osservazione: la formula di Bayes permette di calcolare probabilità dette “a posteriori” (qui significa ad estrazione avvenuta) che indicano “provenienza” in questo caso, ma “colpe”, “meriti”, “responsabilità” ...

Esercizio 2bis: Alle pensiline dei bus Unical

Solo quattro autolinee regionali fanno servizio giornaliero per gli studenti pendolari dell'Unical (Federico, Lirosi, Preite e Romano). Arrivano con frequenze diverse e si può assumere che, su 20 arrivi di autobus al giorno, 8 appartengano a Preite, 6 a Romano, 4 a Lirosi e 2 a Federico.

Si stima che da un (qualsivoglia) bus di Preite scenda alle pensiline il 40% di studenti fuori corso e il 60% di studenti in corso, mentre il 20% di fuori corso scende da un bus di Romano, il 30% da un Lirosi e il 10% da un Federico.

devo prima calcolare la prob totale che sia un fuori corso.
E poi usare sempre Bayes

1. Immaginando di intervistare uno studente pendolare a caso alle pensiline, qual è la probabilità che sia uno dei fuori corso?
2. Accertato che sia fuori corso, qual è la probabilità che provenga dal crotonese, ovvero sia sceso da un bus di Romano?

PROBABILITY AND RANDOM PROCESSES

Venkatarama Krishnan

Professor Emeritus of Electrical Engineering
University of Massachusetts Lowell

Example 2.3.2 Two dice, one red and the other blue, are tossed. These tosses are functionally independent, and we have the Cartesian product of $6 \times 6 = 36$ elementary events in the combined sample space, where each event is equiprobable. We seek the probability that an event B defined by the sum of the numbers showing on the dice equals 9. There are four points $\{(6,3), (5,4), (4,5), (3,6)\}$, and hence $P\{B\} = \frac{4}{36} = \frac{1}{9}$. We now condition the event B with an event A defined as the red die shows odd numbers. The probability of the event A is $P\{A\} = \frac{18}{36} = \frac{1}{2}$. We want to determine whether the events A and B are statistically independent. These events are shown in Fig. 2.3.1, where the first number is for the red die and the second number is for the blue die.

1, 1	1, 2	1, 3	1, 4	1, 5	1, 6
2, 1	2, 2	2, 3	2, 4	2, 5	2, 6
3, 1	3, 2	3, 3	3, 4	3, 5	3, 6
4, 1	4, 2	4, 3	4, 4	4, 5	4, 6
5, 1	5, 2	5, 3	5, 4	5, 5	5, 6
6, 1	6, 2	6, 3	6, 4	6, 5	6, 6

Set A → Set B ←

FIGURE 2.3.1

A 6x6 grid representing a Cartesian product $A \times C$. The vertical axis is labeled $Set A$ and the horizontal axis is labeled $Set C$. The intersection of row 4 and column 5 is shaded.

1, 1	1, 2	1, 3	1, 4	1, 5	1, 6
2, 1	2, 2	2, 3	2, 4	2, 5	2, 6
3, 1	3, 2	3, 3	3, 4	3, 5	3, 6
4, 1	4, 2	4, 3	4, 4	4, 5	4, 6
5, 1	5, 2	5, 3	5, 4	5, 5	5, 6
6, 1	6, 2	6, 3	6, 4	6, 5	6, 6

FIGURE 2.3.2

From Fig. 2.3.1 $P\{A \cap B\} = P\{(3,6), (5,4)\} = \frac{2}{36} = \frac{1}{18}$ and $= P\{A\} \times P\{B\} = \frac{1}{2} \times \frac{1}{9} = \frac{1}{18}$ showing statistical independence. We compute $P\{B | A\}$ from the point of view of reduced sample space. The conditioning event reduces the sample space from 36 points to 18 equiprobable points and the event $\{B | A\} = \{(3,6), (5,4)\}$. Hence $P\{B | A\} = \frac{2}{18} = \frac{1}{9} = P\{B\}$, or, the conditioning event has no influence on B . Here, even though the events A and B are functionally dependent, they are statistically independent.

However, if another set C is defined by the sum being equal to 8 as shown in Fig. 2.3.2, then $P\{C\} = \frac{5}{36}$.

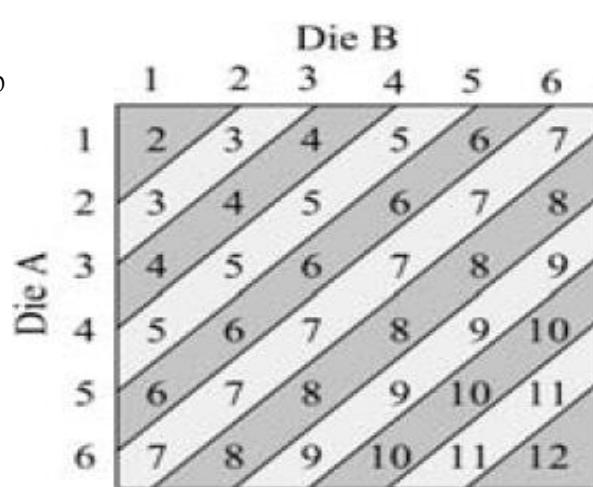
Here the events C and A are not statistically independent because $P\{C\} \cdot P\{A\} = \frac{5}{36} \times \frac{1}{2} = \frac{5}{72} \neq P\{C \cap A\} = P\{(3,5), (5,3)\} = \frac{2}{18} = \frac{4}{72}$.

In this example, we have the case where the events A and C are neither statistically independent nor functionally independent.

Example 2.2.3 The game of craps as played in Las Vegas has the following rules. A player rolls two dice. He wins on the first roll if he throws a 7 or a 11. He loses if the first throw is a 2, 3, or 12. If the first throw is a 4, 5, 6, 8, 9, or 10, it is called a *point* and the game continues. He goes on rolling until he throws the point for a win or a 7 for a loss. We have to find the probability of the player winning.

We will solve this problem both from the definition of conditional probability and the reduced sample space. Figure 2.2.2 shows the number of ways the sums of the pips on the two dice can equal 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, and their probabilities.

da notare che 7 è
il numero con prob
più alta



Total	P{total}
2	$1/36$
3	$2/36$
4	$3/36$
5	$4/36$
6	$5/36$
7	$6/36$
8	$5/36$
9	$4/36$
10	$3/36$
11	$2/36$
12	$1/36$

FIGURE 2.2.2

Solution Using Definition of Conditional Probability. The probability of winning in the first throw is

$$P\{7\} + P\{11\} = \frac{6}{36} + \frac{2}{36} = \frac{8}{36} = 0.22222$$

The probability of losing in the first throw is

$$P\{2\} + P\{3\} + P\{12\} = \frac{1}{36} + \frac{2}{36} + \frac{1}{36} = \frac{4}{36} = 0.11111$$

To calculate the probability of winning in the second throw, we shall assume that i is the point with probability p . The probability of not winning in any given throw after the first is given by $r = P\{\text{not } i \cup \text{ not } 7\} = 1 - p - \frac{1}{6}$. We compute the conditional probability

$$P\{\text{win} \mid i \text{ in first throw}\} = p + rp + r^2p + \dots = \frac{p}{1-r} = \frac{p}{p+\frac{1}{6}}$$

as an infinite geometric series:

$$P(i \text{ in the first row AND win}) = \frac{p^2}{p+\frac{1}{6}} \quad \text{probabilità congiunta}$$

Thus, for $i = 4, 5, 6$, we obtain

$$P\{\text{win after point } i = 4\} = \left(\frac{3}{36}\right)^2 \Big/ \left(\frac{3}{36} + \frac{1}{6}\right) = \frac{1}{36}$$

$$P\{\text{win after point } i = 5\} = \left(\frac{4}{36}\right)^2 \Big/ \left(\frac{4}{36} + \frac{1}{6}\right) = \frac{2}{45}$$

$$P\{\text{win after point } i = 6\} = \left(\frac{5}{36}\right)^2 \Big/ \left(\frac{5}{36} + \frac{1}{6}\right) = \frac{25}{396}$$

with similar probabilities for 8,9,10. Thus the probability of winning in craps is

$$\begin{aligned} P\{\text{win}\} &= P\{\text{win in roll 1}\} + P\{\text{win in roll 2,3,\dots}\} \\ &= \frac{8}{36} + 2 \left[\frac{1}{36} + \frac{2}{45} + \frac{25}{396} \right] = \frac{244}{495} = 0.4929 \end{aligned}$$

Example 2.4.4 (Monty Hall Problem) This classic problem is called the “Monty Hall problem” because of the game show host Monty Hall who designed this game. There are three doors A , B , and C , and behind one of them is a car and behind the other two are goats. The contestant is asked to select any door, and the host Monty Hall opens one of the other two doors and reveals a goat. He then offers the choice to the contestant of switching to the other unopened door or keep the original door. The question now is whether the probability of winning the car is improved if she switches or it is immaterial whether she switches or not.

The answer is counterintuitive in the sense that one may be misled into thinking that it is immaterial whether one switches or not. We will analyze this problem in two different ways.

Mathematical Analysis. Let us analyze this problem from the point of view of Bayes' theorem. We shall first assume that the host has prior knowledge of the door behind which the car is located. Let us also assume that the host opens door B given that the contestant's choice is door A . The a priori probabilities of a car behind the doors A , B , and C are

$$P\{A\} = \frac{1}{3}; \quad P\{B\} = \frac{1}{3}; \quad P\{C\} = \frac{1}{3}$$

We can make the following observations. If the contestant's choice is door A and the car is behind door A , then the host will open the door B with probability of $\frac{1}{2}$ (since she has a choice between B and C). On the other hand, if the car is behind door B , then there is zero probability that she will open door B because of her prior knowledge. If the car is behind door C , then she will open door B with probability 1. Thus we can write the following conditional probabilities for the host opening door B :

$$P\{B | A\} = \frac{1}{2}; \quad P\{B | B\} = 0; \quad P\{B | C\} = 1$$

We can now calculate the total probability $P(B)$ of the host opening door B .

$$\begin{aligned}P\{B\} &= P\{B | A\}P\{A\} + P\{B | B\}P\{B\} + P\{B | C\}P\{C\} \\&= \frac{1}{2} \cdot \frac{1}{3} + 0 \cdot \frac{1}{3} + 1 \cdot \frac{1}{3} = \frac{1}{2}\end{aligned}$$

Using Bayes' theorem [Eq. (2.4.4)] we can now find the a posteriori probabilities of the car behind the doors A or C (B has already been opened by the host) conditioned on B :

$$P\{A | B\} = \frac{P\{BA\}}{P\{B\}} = \frac{P\{B | A\}P\{A\}}{P\{B\}} = \frac{\frac{1}{2} \cdot \frac{1}{3}}{\frac{1}{2}} = \frac{1}{3}$$

$$P\{C | B\} = \frac{P\{BC\}}{P\{B\}} = \frac{P\{B | C\}P\{C\}}{P\{B\}} = \frac{\frac{1}{2} \cdot \frac{1}{3}}{\frac{1}{2}} = \frac{2}{3}$$

Similar analysis holds for other cases of opening the doors A or C . We can see from the result above that the contestant *must* switch if she wants to double her probability of winning the car.

TABLE 2.4.1

First Choice	Host Opens	Second Choice	Probability
A	B	A	$P\{B\} = \frac{1}{2} \cdot \frac{1}{3} = \frac{1}{6}$
A	C	A	$P\{C\} = \frac{1}{2} \cdot \frac{1}{3} = \frac{1}{6}$
B	C	B	$P\{C\} = 1 \cdot \frac{1}{3} = \frac{1}{3}$
C	B	B	$P\{B\} = 1 \cdot \frac{1}{3} = \frac{1}{3}$

TABLE 2.4.2 For not switching

First Choice	Host Opens	Second Choice	Win or Loss	Probability
A	B	A	Win	$P\{B\} = \frac{1}{2} \cdot \frac{1}{3} = \frac{1}{6}$
A	C	A	Win	$P\{C\} = \frac{1}{2} \cdot \frac{1}{3} = \frac{1}{6}$
B	C	B	Loss	$P\{C\} = 1 \cdot \frac{1}{3} = \frac{1}{3}$
C	B	B	Loss	$P\{B\} = 1 \cdot \frac{1}{3} = \frac{1}{3}$

TABLE 2.4.3 For switching

First Choice	Host Opens	Second Choice	Win or Loss	Probability
A	B	C	Loss	$P\{B\} = \frac{1}{2} \cdot \frac{1}{3} = \frac{1}{6}$
A	C	B	Loss	$P\{C\} = \frac{1}{2} \cdot \frac{1}{3} = \frac{1}{6}$
B	C	A	Win	$P\{C\} = 1 \cdot \frac{1}{3} = \frac{1}{3}$
C	B	A	Win	$P\{B\} = 1 \cdot \frac{1}{3} = \frac{1}{3}$

Esercizio: the secretary problem (1/2)

It is one of many names for a famous problem of the [optimal stopping](#) theory. The problem has been studied extensively in the fields of [applied probability](#), and [decision theory](#). It is also known as the **marriage problem**.

1. There is a single secretarial position to fill.
2. There are n applicants for the position, and the value of n is known.
3. The applicants, if seen altogether, can be ranked from best to worst unambiguously.
4. The applicants are interviewed sequentially in random order, with each order being equally likely.
5. Immediately after an interview, the interviewed applicant is either accepted or rejected, and the decision is irrevocable.
6. The decision to accept or reject an applicant can be based only on the relative ranks of the applicants interviewed so far (say “r-1”).
7. The objective of the general solution is to have the highest probability of selecting the best applicant of the whole group.

Esercizio: the secretary problem (2/2)

The optimal policy for the problem is a [stopping rule](#).

Under it, the interviewer rejects the first $r - 1$ applicants (let applicant M be the best applicant among these $r - 1$ applicants), and then selects the first subsequent applicant that is better than applicant M .

It can be shown that the optimal strategy lies in this class of strategies.

For an arbitrary cutoff r , the probability that the best applicant is selected is:

$$\begin{aligned} P(r) &= \sum_{i=1}^n P(\text{applicant } i \text{ is selected} \cap \text{applicant } i \text{ is the best}) && \text{prova a farlo se riesci con} \\ &= \sum_{i=1}^n P(\text{applicant } i \text{ is selected} | \text{applicant } i \text{ is the best}) \times P(\text{applicant } i \text{ is the best}) && \text{probabilità condizionata ecc} \\ &= \left[\sum_{i=1}^{r-1} 0 + \sum_{i=r}^n P\left(\begin{array}{l} \text{the second best of the first } i \text{ applicants} \\ \text{is in the first } r-1 \text{ applicants} \end{array} \middle| \text{applicant } i \text{ is the best} \right) \right] \times \frac{1}{n} \\ &= \sum_{i=r}^n \frac{r-1}{i-1} \times \frac{1}{n} = \frac{r-1}{n} \sum_{i=r}^n \frac{1}{i-1}. \end{aligned}$$



n	1	2	3	4	5	6	7	8	9
r	1	1	2	2	3	3	3	4	4
P	1.000	0.500	0.500	0.458	0.433	0.428	0.414	0.410	0.406

EXTRA

Esercizi di Riepilogo

Esercizio 1

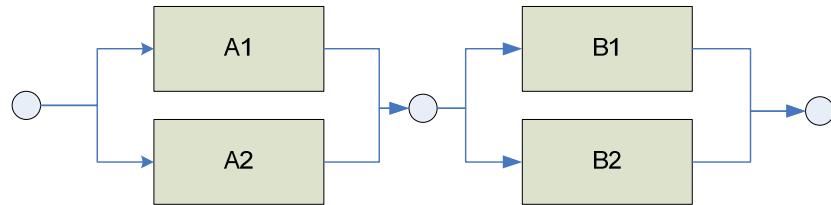
Affinché il sistema dei 4 motori che compongono un grande aereo di linea sia considerato funzionante è richiesto che risultino funzionanti almeno 2 motori su 4 e che i due funzionanti non stiano entrambi sotto la stessa ala. Intesa come probabilità di successo, l'affidabilità di ciascuno dei 4 motori (identici) può essere considerata pari a 9×10^{-4} .

- i.) Presentare le assunzioni che si è costretti ad introdurre per poter modellare l'affidabilità del componente/sistema come una “probabilità di successo”;
- ii.) Disegnare il diagramma a blocchi di affidabilità evidenziandone i cammini;
- iii.) Assumendo che il guasto di un motore non influisca sulla probabilità di guasto di alcun altro motore, calcolare l'affidabilità del sistema, anche ragionando con i cammini individuabili;
- iv.) Assumendo, invece, che l'evento di guasto di un motore influisca sulla probabilità di guasto del motore adiacente (sotto la stessa ala), perché entrambi condividono soltanto un terminale di alimentazione che può risultare guasto con probabilità 8×10^{-5} , ricalcolare l'affidabilità di un motore;
- v.) Rilassando il vincolo che “debba funzionare almeno un motore sotto ciascuna ala”, ricalcolare l'affidabilità del sistema, valutando l'approssimazione rispetto al valore vero ottenuto al punto iii.

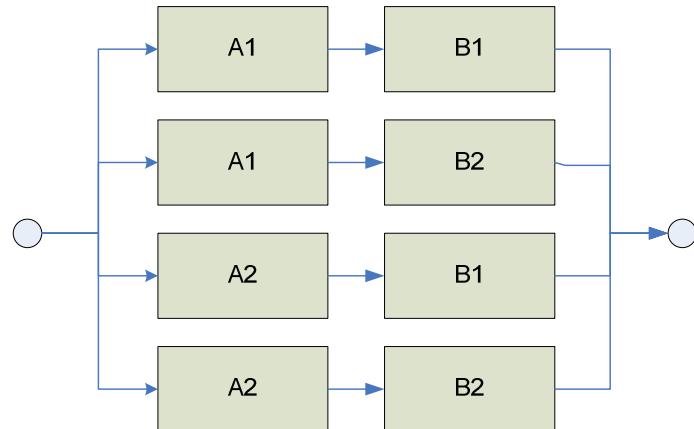
Soluzione

- i.) Le assunzioni che si è costretti ad introdurre per poter modellare l'affidabilità del componente/sistema come una “probabilità di successo” sono così riassumibili:
 - è possibile ispezionare lo stato di funzionamento di ciascun componente in tempi trascurabili;
 - l'istante di tempo in cui avviene l'ispezione non è importante;
 - l'affidabilità è esprimibile come una grandezza binaria: 1 sistema funzionante, 0 sistema non funzionante;
 - il concetto di durata non è contemplata nell'affidabilità intesa come “probabilità di successo”.

ii.) Il diagramma a blocchi del sistema è:



Il diagramma a blocchi di affidabilità basato sui cammini è il seguente:



iii.) Assumendo che il guasto di un motore non influisca sulla probabilità di guasto di alcun altro motore, calcolare l'affidabilità del sistema, anche ragionando con i cammini individuabili;

Si introducono le seguenti notazioni:

$A \hat{=} \text{il sistema funziona}$	$R_j \text{ è l'affidabilità del componente } j$	$C_i \text{ è il cammino } i$
	componente 1 = A_1	$C1 \hat{=} A_1 \cap B_1$
	componente 2 = A_2	$C2 \hat{=} A_1 \cap B_2$
	componente 3 = B_1	$C3 \hat{=} A_2 \cap B_1$
	componente 4 = B_2	$C4 \hat{=} A_2 \cap B_2$

Affidabilità del sistema calcolato come il seriale di due paralleli è:

$$R_{P1} = 1 - (1 - R_1)(1 - R_2) = 1 - (1 - 9 \cdot 10^{-4})(1 - 9 \cdot 10^{-4}) \cong 1,8 \cdot 10^{-3}$$

$$R_{P2} = 1 - (1 - R_3)(1 - R_4) = 1 - (1 - 9 \cdot 10^{-4})(1 - 9 \cdot 10^{-4}) \cong 1,8 \cdot 10^{-3}$$

$$R_{SP} = R_{P1} \cdot R_{P2} = 3,24 \cdot 10^{-6}$$

Affidabilità del sistema calcolato a partire dai cammini individuali:

$$P(A) = P\left(\bigcup_{i=1}^4 C_i\right) = P(C_1 \cup C_2 \cup C_3 \cup C_4)$$

$$\begin{aligned}
&= P(C_1) + P(C_2) + P(C_3) + P(C_4) - P(C_1 \cap C_2) - P(C_1 \cap C_3) - P(C_1 \cap C_4) - P(C_2 \cap C_3) + \\
&\quad - P(C_2 \cap C_4) - P(C_3 \cap C_4) + P(C_1 \cap C_2 \cap C_3) + P(C_1 \cap C_2 \cap C_4) + P(C_1 \cap C_3 \cap C_4) + \\
&\quad + P(C_2 \cap C_3 \cap C_4) - P(C_1 \cap C_2 \cap C_3 \cap C_4) \\
&= P(C_1) + P(C_2) + P(C_3) + P(C_4) - P(C_1 | C_2) \cdot P(C_2) - P(C_1 | C_3) \cdot P(C_3) + \\
&\quad - P(C_1 | C_4) \cdot P(C_4) - P(C_2 | C_3) \cdot P(C_3) - P(C_2 | C_4) \cdot P(C_4) - P(C_3 | C_4) \cdot P(C_4) + \\
&\quad + P(C_1 | C_2 \cap C_3) \cdot P(C_2 \cap C_3) + P(C_1 | C_2 \cap C_4) \cdot P(C_2 \cap C_4) + \\
&\quad + P(C_1 | C_3 \cap C_4) \cdot P(C_3 \cap C_4) + P(C_2 | C_3 \cap C_4) \cdot P(C_3 \cap C_4) + \\
&\quad - P(C_1 | C_2 \cap C_3 \cap C_4) \cdot P(C_2 \cap C_3 \cap C_4) \\
&= P(C_1) + P(C_2) + P(C_3) + P(C_4) - P(C_1 | C_2) \cdot P(C_2) - P(C_1 | C_3) \cdot P(C_3) + \\
&\quad - P(C_1 | C_4) \cdot P(C_4) - P(C_2 | C_3) \cdot P(C_3) - P(C_2 | C_4) \cdot P(C_4) - P(C_3 | C_4) \cdot P(C_4) + \\
&\quad + P(C_1 | C_2 \cap C_3) \cdot P(C_2 | C_3) \cdot P(C_3) + P(C_1 | C_2 \cap C_4) \cdot P(C_2 | C_4) \cdot P(C_4) + \\
&\quad + P(C_1 | C_3 \cap C_4) \cdot P(C_3 | C_4) \cdot P(C_4) + P(C_2 | C_3 \cap C_4) \cdot P(C_3 | C_4) \cdot P(C_4) + \\
&\quad - P(C_1 | C_2 \cap C_3 \cap C_4) \cdot P(C_2 | C_3 \cap C_4) \cdot P(C_3 | C_4) \cdot P(C_4) \\
&= (R_1 \cdot R_3) + (R_1 \cdot R_4) + (R_2 \cdot R_3) + (R_2 \cdot R_4) - (R_1 \cdot R_3 \cdot R_4) - (R_1 \cdot R_3 \cdot R_2) + \\
&\quad - (R_1 \cdot R_3 \cdot R_2 \cdot R_4) - (R_1 \cdot R_4 \cdot R_2 \cdot R_3) - (R_1 \cdot R_2 \cdot R_4) - (R_3 \cdot R_2 \cdot R_4) + \\
&\quad + (R_1 \cdot R_4 \cdot R_2 \cdot R_3) + (R_3 \cdot R_1 \cdot R_2 \cdot R_4) + (R_1 \cdot R_3 \cdot R_2 \cdot R_4) + (R_1 \cdot R_3 \cdot R_2 \cdot R_4) + \\
&\quad - (R_1 \cdot R_3 \cdot R_2 \cdot R_4) \\
&= 81 \cdot 10^{-8} + 81 \cdot 10^{-8} + 81 \cdot 10^{-8} + 81 \cdot 10^{-8} - 7,29 \cdot 10^{-10} - 7,29 \cdot 10^{-10} - 6,56 \cdot 10^{-13} + \\
&\quad - 6,56 \cdot 10^{-13} - 7,29 \cdot 10^{-10} - 7,29 \cdot 10^{-10} + 6,56 \cdot 10^{-13} + 6,56 \cdot 10^{-13} + 6,56 \cdot 10^{-13} + \\
&\quad + 6,56 \cdot 10^{-13} - 6,56 \cdot 10^{-13} \cong 3,24 \cdot 10^{-6}
\end{aligned}$$

iv.) Assumendo, invece, che l'evento di guasto di un motore influisca sulla probabilità di guasto del motore adiacente (sotto la stessa ala), perchè entrambi condividono

soltanto un terminale di alimentazione che può risultare guasto con probabilità $8 \cdot 10^{-5}$, ricalcolare l'affidabilità di un motore;

$\bar{A}_j \hat{=} \text{il componente } j \text{ non funziona}$

$\bar{A}_1 \cap \bar{A}_2 = 8 \cdot 10^{-5}$ è la probabilità di guasto comune tra i due componenti sotto la medesima ala.

Applicando la formula della probabilità totale, si ottiene la nuova probabilità che il componente A_1 si guasti ovvero \bar{A}_1

$$\begin{aligned} P^{new}(\bar{A}_1) &= P(\bar{A}_1 | A_2) \cdot P(A_2) + P(\bar{A}_1 | \bar{A}_2) \cdot P(\bar{A}_2) \\ &= P(\bar{A}_1) \cdot P(A_2) + P(\bar{A}_1 | \bar{A}_2) \cdot P(\bar{A}_2) \end{aligned}$$

dove

$$P(\bar{A}_1 | \bar{A}_2) = P(\bar{A}_1 \cap \bar{A}_2) / P(\bar{A}_2) = 8 \cdot 10^{-5} / (1 - 9 \cdot 10^{-4}) = 8,00721 \cdot 10^{-5}.$$

Segue

$$R_{A_1}^{new} = 1 - P^{new}(\bar{A}_1).$$

In numeri

$$P^{new}(\bar{A}_1) = (1 - 9 \cdot 10^{-4}) \cdot (9 \cdot 10^{-4}) + (8,00721 \cdot 10^{-5}) \cdot (1 - 9 \cdot 10^{-4}) = 9,79 \cdot 10^{-4}$$

da cui

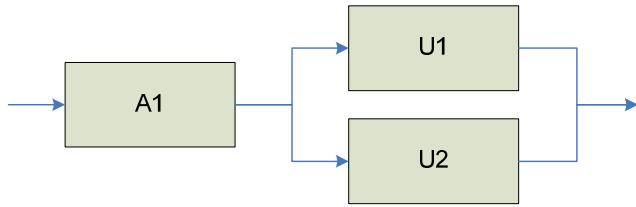
$$R_{A_1}^{new} = 1 - P^{new}(\bar{A}_1) = 1 - 9,79 \cdot 10^{-4} = 0,999021$$

v.) Rilassando il vincolo che “debbra funzionare almeno un motore sotto ciascuna ala”, ricalcolare l'affidabilità del sistema, valutando l'approssimazione rispetto al valore vero ottenuto al punto iii.

Rilassare il suddetto vincolo significa che almeno due motori su quattro devono funzionare, a prescindere dall'ala sotto la quale sono posizionati. Ciò premesso, la nuova affidabilità del sistema è pari a:

$$\begin{aligned} R_{\text{sistema}} &= \sum_{j=2}^4 \binom{4}{i} R^j \cdot (1-R)^{4-j} \\ &= 6 R^2 \cdot (1-R)^{4-2} + 4 R^3 \cdot (1-R)^{4-3} + R^4 \cdot (1-R)^{4-4} \cong 4,85 \cdot 10^{-6} \end{aligned}$$

Come lecito aspettarsi, tale affidabilità costituisce un upper-bound rispetto al valore calcolato al punto iii ($4,85 \cdot 10^{-6} \geq 3,24 \cdot 10^{-6}$), superiore del 67% circa.



3. Posto che il sistema è guasto, la probabilità che detto guasto sia stato causato solo dal componente A2 fornisce un indice di criticità del componente.

$$\begin{aligned}
 P(\bar{A}_2 | \bar{A}) &= P(\bar{A} | \bar{A}_2) \cdot P(\bar{A}_2) / P(\bar{A}) \\
 &= \{1 - R_{A_1} \cdot [1 - (1 - R_{U_1}) \cdot (1 - R_{U_2})]\} \cdot (1 - R_{A_2}) / P(\bar{A}) \\
 &= \{1 - 0,98 \cdot [1 - (1 - 0,9) \cdot (1 - 0,96)]\} \cdot (1 - 0,98) / (1 - 0,999) \approx 0,4784
 \end{aligned}$$

Esercizio 3

Un sistema di produzione è caratterizzato da quattro “isole” dove vengono realizzati componenti meccanici pesanti, che dovranno essere trasferiti ad una ulteriore isola di assemblaggio e collaudo del prodotto finito. Il trasporto avviene mediante una navetta elettrica, che può accogliere un solo componente per volta. La navetta è disponibile per servire una qualunque delle quattro isole soltanto in corrispondenza di istanti che si ripetono con cadenza fissa. In corrispondenza di uno qualunque di tali istanti, il sistema di controllo dell’impianto di produzione rileva l’eventuale richiesta avanzata da una o più delle suddette isole di produzione ed invia la navetta ad una di esse.

Tenendo presente quanto segue:

- La probabilità che in uno stesso istante si verifichino proprio 0, 1, 2, 3 o 4 richieste è pari, rispettivamente, a: 0,30, 0,50, 0,10, 0,06, 0,04;
- Nel caso in cui vengano rilevate 2 o più richieste di trasporto, il sistema di controllo risolve il conflitto scegliendo in maniera equiprobabile l’isola alla quale inviare la navetta;
- Un’isola che non vede soddisfatta la propria richiesta di trasporto all’istante in cui l’ha fatta, rinnoverà la richiesta in un istante successivo, scelto col meccanismo del “lancio della monetina”, oppure rinuncerà del tutto ad usare la navetta per quello specifico componente, optando per una modalità alternativa;

si chiede di calcolare:

1. la probabilità che ci sia conflitto per l'assegnazione della navetta;
2. la probabilità che un'isola di produzione possa ottenere la navetta proprio nell'istante in cui ne fa richiesta;
3. la probabilità che un'isola di produzione si veda assegnata la navetta proprio alla terza richiesta consecutiva;
4. la probabilità che un'isola di produzione debba avanzare al più due richieste per vedersi assegnata la navetta;
5. la probabilità che un'isola di produzione debba fare più di 3 richieste per riuscire ad utilizzare la navetta 2 volte, non necessariamente consecutive.

Soluzione

1. Una situazione di conflitto per l'assegnazione della navetta si ha quando due o più isole di produzione avanzano richiesta di trasporto

Se $n \hat{=} \text{numero di richieste di trasporto}$ allora

$$\begin{aligned} P(n \geq 2) &= P(2) + P(3) + P(4) \\ &= 0,1 + 0,06 + 0,04 = 0,2 \end{aligned}$$

2. Per calcolare la probabilità che un'isola di produzione possa ottenere la navetta proprio nell'istante in cui ne fa richiesta si ricorre alla formula della probabilità totale. Ma prima, per ottenere una partizione dello spazio dei risultati, occorre ridefinire le probabilità in quanto il quesito presume che almeno una isola faccia richiesta di trasporto. Quindi, occorre escludere la probabilità che in uno stesso istante non si verifichi alcuna richiesta.

$$\begin{aligned} P^{new}(1) &= P(1)/(P(1) + P(2) + P(3) + P(4)) \\ &= 0,5 / 0,7 = 0,714 \end{aligned}$$

$$\begin{aligned} P^{new}(2) &= P(2)/(P(1) + P(2) + P(3) + P(4)) \\ &= 0,1 / 0,7 = 0,143 \end{aligned}$$

$$P^{new}(3) = P(3)/(P(1)+P(2)+P(3)+P(4)) \\ = 0,06/0,7 = 0,086$$

$$P^{new}(4) = P(4)/(P(1)+P(2)+P(3)+P(4)) \\ = 0,04/0,7 = 0,057$$

Alla luce della suddetta ridefinizione, la probabilità che un'isola di produzione possa ottenere la navetta proprio nell'istante in cui ne fa richiesta è pari a:

$$P(\text{ottenere la navetta}) = P(\text{ottenere la navetta} | 1 \text{ richiesta}) \cdot P^{new}(1 \text{ richiesta}) + \\ + P(\text{ottenere la navetta} | 2 \text{ richieste}) \cdot P^{new}(2 \text{ richieste}) + \\ + P(\text{ottenere la navetta} | 3 \text{ richieste}) \cdot P^{new}(3 \text{ richieste}) + \\ + P(\text{ottenere la navetta} | 4 \text{ richieste}) \cdot P^{new}(4 \text{ richieste}) + \\ = 1 \cdot 0,714 + 1/2 \cdot 0,143 + 1/3 \cdot 0,143 + 1/4 \cdot 0,057 = 0,828$$

3. La probabilità che un'isola di produzione si veda assegnata la navetta proprio alla terza richiesta consecutiva si calcola come caso particolare del modello di Bernoulli secondo il quale la probabilità di ottenere il primo successo al tentativo "k-simo"

$$P(\text{primo successo alla prova } k) = q^{k-1} p \text{ con } p = 0,828, q = 1 - p = 0,172 \text{ e } k = 3 \\ = 0,172^2 \cdot 0,828 \cong 0,024$$

4. La probabilità che un'isola di produzione debba avanzare al più due richieste per vedersi assegnata la navetta è ottenuta applicando ancora la formula di cui al punto precedente nei casi in cui $1 \leq k \leq 2$.

$$P(\text{successo con al più } k \text{ prove}) = \sum_{k=1}^2 q^{k-1} p \text{ con } p = 0,828 \text{ e } q = 1 - p = 0,172 \\ = 0,828 + 0,172 \cdot 0,828 \cong 0,970$$

5. La probabilità che un'isola di produzione debba fare più di 3 richieste per riuscire ad utilizzare la navetta 2 volte, non necessariamente consecutive, si calcola a partire dalla binomiale negativa.

$$\begin{aligned}
 P(n \text{ prove per ottenere } k \text{ successi}) &= \sum_{n=2}^3 \binom{n-1}{k-1} p^k q^{n-k} \text{ con } p = 0,828, q = 1-p = 0,172 \text{ e } k = 2 \\
 &= 0,828^2 \cdot 1 + 2 \cdot 0,828^2 \cdot 0,172 \cong 0,921
 \end{aligned}$$

Questa esprime la probabilità che occorrono n prove (con $2 \leq n \leq 3$) per ottenere un numero di successi pari a k (con $k = 2$). Quindi, per ottenere la probabilità che si debbano fare più di 3 richieste è sufficiente fare il complemento ad 1 di $P(n \text{ prove per ottenere } k \text{ successi})$, cioè

$$\begin{aligned}
 P(\text{più di } n \text{ prove per } k \text{ successi}) &= 1 - P(n \text{ prove per ottenere } k \text{ successi}) \\
 &= 1 - 0,921 = 0,079
 \end{aligned}$$

Esercizio 4

I valori di pressione e temperatura in un preciso punto critico di un impianto a fluido sono controllati continuamente da un sistema elettronico composto da alcuni sensori, una logica di controllo ed alcuni attuatori. La scatola di controllo è stata replicata in 4 moduli identici che elaborano gli stessi dati prelevati dai sensori; le azioni di regolazione dei parametri di temperatura e pressione dell'impianto vengono realizzate dagli attuatori, ma solo se il comparatore posto a valle dei 4 moduli di controllo rileva un accordo nella misura di almeno 3 su quattro.

Sapendo che la probabilità di errore di ognuno dei 4 moduli di controllo è pari a 0.05, si chiede di calcolare:

1. l'affidabilità del sistema di controllo;
2. la probabilità che i 4 moduli diano tutti risultati errati;
3. la probabilità che il sistema funzioni al minimo;
4. la probabilità che funzioni al massimo delle sue prestazioni, posto che funzioni;

5. la probabilità che funzioni il modulo numero 1 ed almeno 2 dei tre rimanenti.

Soluzione

1. Il sistema di controllo è affidabile se almeno tre su quattro dei moduli identici funzionano correttamente.

Se si indica con $R_{sistema} \hat{=} \text{Affidabilità del sistema}$ e con $R \hat{=} \text{Affidabilità del singolo modulo} = 0,95$, allora l'affidabilità di sistema si ricava impiegando la formula per il calcolo dell'affidabilità di un sistema “m-out-of-n” (in questo caso, 3 moduli su 4).

$$\begin{aligned} R_{sistema} &= \sum_{i=3}^4 \binom{4}{i} R^i \cdot (1-R)^{4-i} \\ &= 4 \cdot 0,95^3 \cdot (1-0,95) + 0,95^4 \cong 0,986 \end{aligned}$$

2. La probabilità che i 4 moduli diano tutti risultati errati è data dal complemento ad 1 che i 4 moduli siano tutti funzionanti.

$$\begin{aligned} P(\text{nessun modulo funzionante}) &= 1 - P(\text{esattamente 4 moduli funzionanti}) \\ &= 1 - \sum_{i=4}^4 \binom{4}{i} R^i \cdot (1-R)^{4-i} = 1 - R^4 \\ &= 1 - 0,95^4 \cong 0,815 \end{aligned}$$

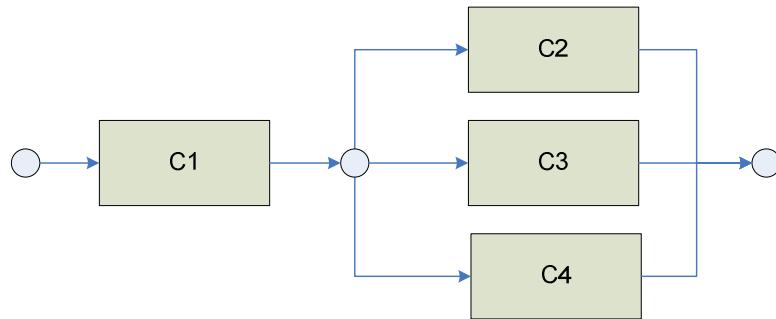
3. La probabilità che il sistema funzioni al minimo si ha quando esattamente 3 moduli funzionano.

$$\begin{aligned} P(\text{esattamente 3 moduli funzionanti}) &= \binom{4}{3} R^3 \cdot (1-R) \\ &= 4 \cdot 0,95^3 \cdot (1-0,95) \cong 0,172 \end{aligned}$$

4. La probabilità che il sistema funzioni al massimo delle sue prestazioni si ha quando, posto che il sistema funzioni, esattamente 4 moduli funzionano.

$$\begin{aligned}
 P(\text{tutti moduli funzionanti} \mid \text{sistema funziona}) &= \frac{P(4 \cap R_{\text{sistema}})}{P(R_{\text{sistema}})} \\
 &= \frac{0,95^4}{0,986} \cong 0,826
 \end{aligned}$$

5. La probabilità che funzioni il modulo numero 1 ed almeno 2 dei tre rimanenti è calcolato a partire dalla seguente rappresentazione.



L'affidabilità di sistema è dato dall'affidabilità della serie del componente C_1 ed il sistema m-out-of-n a valle (con $n = 3$ e $m = 2$).

$$\begin{aligned}
 R_{\text{sistema}} &= R_{C1} \cdot R_{2-\text{out}-\text{of}-3} \\
 &= R \cdot \left[\sum_{i=2}^3 \binom{3}{i} R^i \cdot (1-R)^{n-i} \right] \\
 &= 0,95 \cdot [3 \cdot 0,95^2 \cdot (1-0,95) + 0,95^3] \cong 0,943
 \end{aligned}$$

SECONDA PARTE ANALISI PROBABILISTA

Variabili aleatorie e distribuzioni

Dato uno spazio di risultati (anche non numerico), Ω di un esperimento aleatorio una variabile aleatoria X è una funzione che associa ad ognuno dei risultati ω numeri reali. Per ipotesi i risultati dell'esperimento saranno numeri. Preso un insieme di risultati, se si studia il suo andamento temporale andremo a trattare di un processo stocastico.

Variabile aleatoria: $X : \Omega \rightarrow \mathbb{R}$ Processo stocastico: $X_\omega : \mathbb{R} \rightarrow S$ con $\omega \in \Omega, S \subseteq \mathbb{R}$

Una volta eseguito l'esperimento, il valore $X(\omega) = x$ risulta completamente determinato e viene detto **realizzazione** (come definito anche dopo)

Fortunatamente, in questa sede, è sufficiente concentrarsi su risultati che sono già

- 1) numeri interi non negativi: come nel caso di "k successi su n prove" oppure nel caso di "esattamente/al più/almeno k posti occupati in un buffer";
- 2) numeri reali non negativi: come nel caso della "durata del servizio reso da una risorsa ad un utente", oppure della "durata dell'attesa dell'utente prima di ottenere la risorsa", oppure del "tempo al guasto di un componente/sistema".

Insomma si riduce di molto lo spazio dei risultati

Caso 1) → variabili aleatorie discrete (a valori nel discreto).

Caso 2) → variabili aleatorie continue (a valori nel continuo).

Nel caso 2) saranno considerate variabili aleatorie che fanno corrispondere il numero reale x - detto **realizzazione** della variabile aleatoria X - all'evento "la durata del servizio/attesa (o del fenomeno qualsivoglia) è non superiore a x ". Dunque tutti i possibili risultati nel continuo (precise durate del fenomeno) compresi nell'intervallo reale $[0,x]$ verranno posti in corrispondenza alla realizzazione x della X e ciò verrà indicato con la notazione " $X \leq x$ ", intendendo che la X è la variabile aleatoria che rappresenta l'insieme delle possibili durate del fenomeno. L'insieme dei numeri (ovvero le durate) sarà considerato positivo, cioè da 0 a infinito perché durate negative non hanno senso

Tornando al caso discreto, si consideri la variabile aleatoria, K , che rappresenta "il numero di possibili successi su n prove di Bernoulli". Una sua realizzazione, k , corrisponde all'evento "k successi su n prove", del quale è stata già calcolata la probabilità:

$$\binom{n}{k} p^k q^{n-k}$$

È allora naturale definire distribuzione di probabilità della variabile aleatoria discreta K la funzione $P : [0,1,2,\dots,n] \rightarrow [0,1] \in \Re$ che ripartisce i singoli valori di probabilità alle singole realizzazioni, $k=0,1,\dots,n$, della K rispettando gli assiomi della probabilità.

$$P(K = k | n) \equiv P_K(k | n) \triangleq \binom{n}{k} p^k q^{n-k}, \quad k = 0,1,\dots,n;$$

La funzione appena definita (nota come *legge binomiale*, $B(n,p)$) è un assegnamento probabilistico sullo spazio delle realizzazioni della K perché risulta:

- 1) $0 \leq \binom{n}{k} p^k q^{n-k} \leq 1 \quad \forall k = 0,1,\dots,n$ dunque probabilità compresa tra 0 e 1
- 2) $\sum_{k=0}^n \binom{n}{k} p^k q^{n-k} = (p+q)^n = 1 \quad$ sarebbe il binomio di Newton
appunto somma delle probabilità pari a 1

Ad un'altra legge di probabilità ben nota, *la legge geometrica*, si può pervenire introducendo la variabile aleatoria, S, che rappresenta "l'insieme dei numeri di prova in corrispondenza dei quali si può ottenere il primo successo, nell'ambito di una successione indefinita di prove di Bernoulli"

Infatti, partendo dall'evento "il primo successo al k-mo tentativo"
si pone:

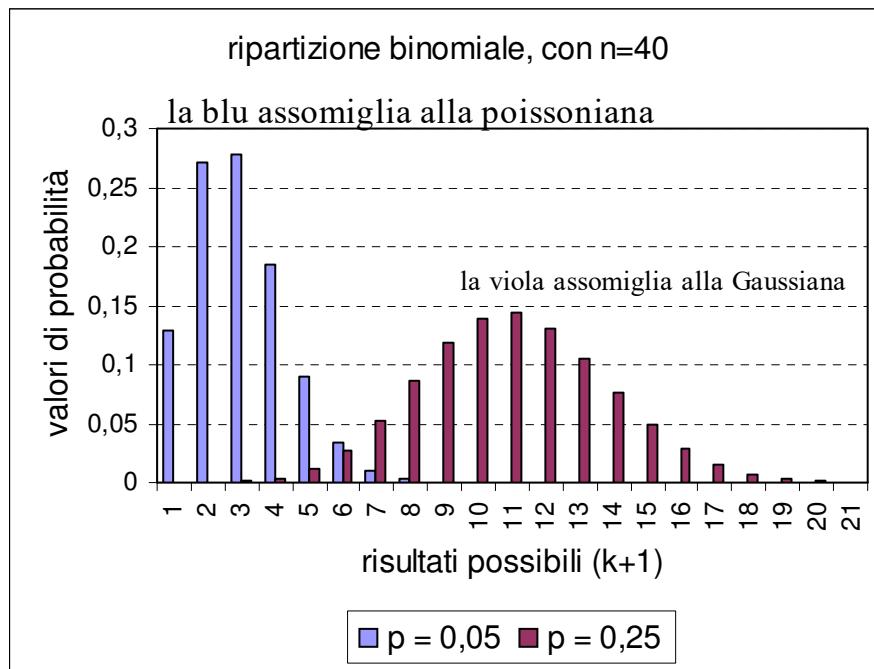
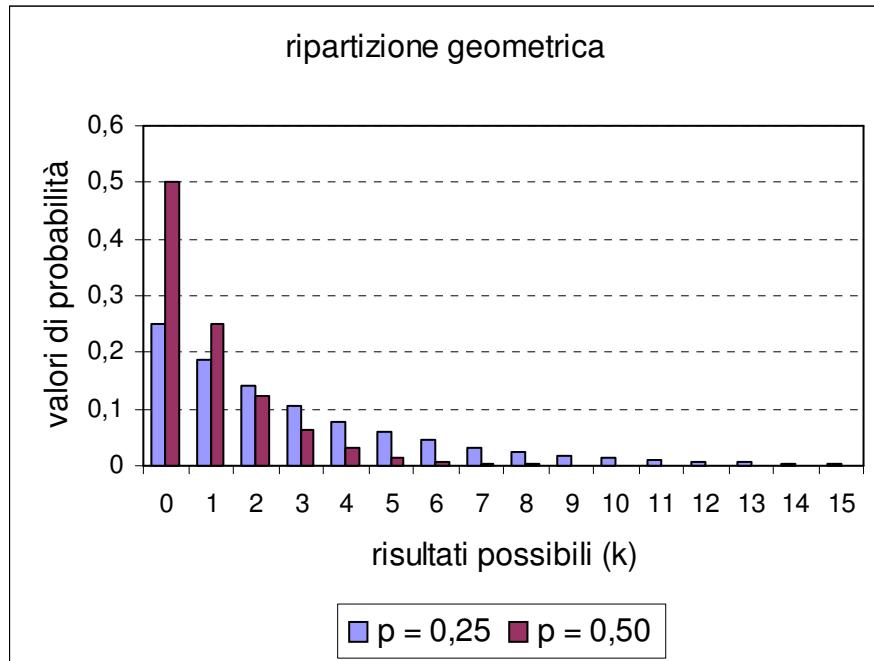
$$P(S = k | \infty) \equiv P_S(k | \infty) \triangleq q^{k-1} p, \quad k = 1, \dots, n, \dots \quad \begin{array}{l} \text{si parte da 1 perché deve} \\ \text{accadere sto successo} \end{array}$$

e si verifica che:

$$0 \leq q^{k-1} p \leq 1 \quad \forall k = 1, \dots, n, \dots \quad \text{e} \quad \sum_{k=1}^{\infty} q^{k-1} p = \dots = p/(1-q) = 1$$

anche qui la probabilità è tra 0 e 1 e la somma di tutte le probabilità ritorna 1

Rappresentazioni di Geometrica e Binomiale



Esempio di costruzione di un modello probabilistico di ripartizione

A partire dalla seguente (lo sviluppo in serie della funzione esponenziale positiva)

$$\sum_{n=0}^{\infty} \frac{x^n}{n!} = e^x \quad x > 0$$

E moltiplicando ambo i membri per e^{-x} si ottiene

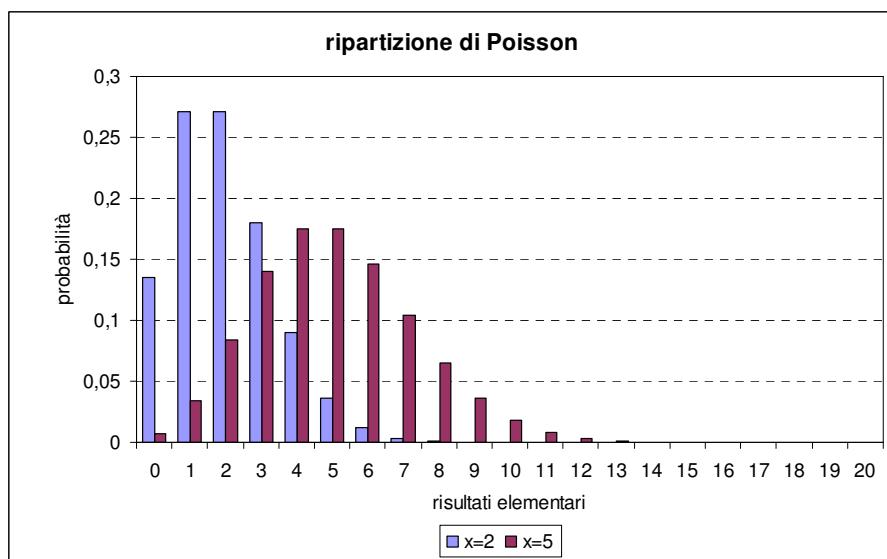
$$\sum_{n=0}^{\infty} \frac{x^n}{n!} e^{-x} = 1 \quad x > 0$$

Osservando che il termine dentro la sommatoria $\frac{x^n}{n!} e^{-x} \geq 0$ rispetta l'assioma (A1) della probabilità (di non negatività), mentre la sommatoria uguale ad 1 può essere letta come la verifica dell'assioma (A3), è immediato concludere che la funzione che scriviamo

$$P(N = n) = \frac{x^n}{n!} e^{-x}$$

dove $n = 0, 1, 2, \dots$ sono i risultati possibili della variabile aleatoria N ed x è un parametro di detta funzione.

La funzione è un nuovo modello di ripartizione (noto come modello di Poisson).



IMPORTANTE
Distribuzione(cumulativa)di probabilità

Per introdurre il concetto e la definizione di distribuzione cumulativa di probabilità per una variabile aleatoria discreta, è utile fare riferimento all'esperimento aleatorio che consiste nell'osservazione di un buffer e al risultato "*al più i primi m posti risultano occupati*", sotto l'ipotesi che il numero di posti disponibili sia praticamente illimitato.

Introducendo la variabile aleatoria K per rappresentare tutti i risultati (elementari) del tipo "*i primi k posti risultano occupati*", con $k = 0, 1, 2, \dots, n, \dots$, si potrebbe adottare la legge geometrica per ripartire valori di probabilità alle possibili realizzazioni:

$$P(K = k | \infty) \equiv P_K(k | \infty) \hat{=} q^k p, \quad k = 0, 1, \dots, n, \dots$$

assumendo che il fattore "p" sia la probabilità di trovare un posto libero dopo k posti trovati tutti occupati (ognuno con la stessa probabilità "q").

A questo punto, per calcolare la probabilità del risultato d'interesse, "*al più i primi m posti risultano occupati*", basta cumulare (sommare) le probabilità ripartite dalla legge geometrica su tutti i casi elementari compresi fra "0" ed "m".

Dunque:

$$\text{Probabilità("al più i primi m posti occupati")} = \sum_{k=0}^m q^k p \hat{=} C(m)$$

L'importanza della probabilità appena calcolata giustifica abbastanza la definizione della funzione C , sul dominio dei possibili valori di m , quale distribuzione cumulativa associata ad una data (qui era la geometrica) distribuzione di probabilità nel discreto.

Affinché la C sia ben posta deve sempre risultare:

$$0 \leq C(m) \leq 1 \quad \forall m \quad \text{e} \quad C(m) \rightarrow 1 \quad \text{se } m \rightarrow \infty$$

$C(m)$ resta una probabilità!

IMPORTANTE

Funzione di distribuzione

Con riferimento al caso continuo, l'utilità di misurare la probabilità che "la durata di un certo fenomeno sia non superiore ad un valore reale positivo, x , fissato" conduce alla definizione di funzione di distribuzione associata alla variabile continua X :

$$F_X : [0 \leq x < \infty] \rightarrow [0,1] \in \Re \quad \begin{array}{l} \text{codominio da 0 a 1} \\ \text{perché è una probabilità} \end{array}$$

La probabilità che la durata sia al più pari a x è pari a

$$P(X \leq x) \equiv F_X(x) \hat{=} \int_{t=0}^x f_X(t) dt \quad \begin{array}{l} \text{l'estremo superiore è } x, \text{ dunque} \\ f_X \text{ è una funzione integrale} \end{array}$$

Proprietà di F_X

l'integrale sarebbe la somma delle probabilità

$$0 \leq F_X(x) \leq 1, \quad 0 \leq x < \infty$$

Banalmente la cumulativa deve essere nulla all'inizio perché non hai accumulato nulla!

$$F_X(x_2) \geq F_X(x_1), \quad x_2 \geq x_1 \quad \begin{array}{l} \text{è monotona} \\ \text{se } x=0 \text{ allora la funzione è pari a 0} \end{array}$$

"alla fine hai terminato"

$$F_X(x \rightarrow 0) \rightarrow 0, \quad F_X(x \rightarrow \infty) \rightarrow 1 \quad \begin{array}{l} \text{ovvero probabilità è} \\ \text{pari 1} \end{array}$$

La f_X è detta funzione densità di probabilità ed ha il compito di ripartire valori di probabilità alle realizzazioni, continue e non negative, della X .

Dunque detta in parole povere, densità di probabilità significa: quanta probabilità c'è in punto? Ad esempio 0.4 a $t=1$, 1.0 a $t=0$.

A livello elementare, si può assumere che la densità di probabilità sia una funzione continua e quindi integrabile secondo Jordan, nonché ottenibile per derivazione in base alla seguente $f_X : [0 \leq x < \infty] \rightarrow [0, M] \in \Re^+, \quad M > 1$

M può essere molto maggiore di 1
 $f_X(x)dx$ è una probabilità
 che il fenomeno duri proprio x a
 meno di un dx . Esempio probabilità
 che un download arriva a 30s
 e finisce tra 30s e 31s. Quindi
 $dx = 1s$.

$$0 \leq f_X(x) \mid \int_{x=0}^{\infty} f_X(x) dx = 1$$

la derivata si usa perché non si può sapere l'istante esatto

$$\frac{d}{dx} F_X(x) = \frac{d}{dx} \left(\int_{t=0}^x f_X(t) dt \right) = f_X(x)$$

in pratica dalla distribuzione
 passo alla densità tramite derivata

la derivata sarebbe il teorema fondamentale

$$P(x_1 \leq X \leq x_2) = \int_{t=x_1}^{t=x_2} f_X(t) dt, \quad \geq 0 = F(x_2)$$

data la probabilità tra x_1 e x_2

$$P(x \leq X \leq x + \Delta x) = F_X(x + \Delta x) - F_X(x)$$

$$\cong f_X(x)\Delta x, \quad \Rightarrow P(X = x) = 0 \quad \begin{array}{l} \text{risulta nulla la probabilità} \\ \text{associabile ad un punto} \end{array}$$

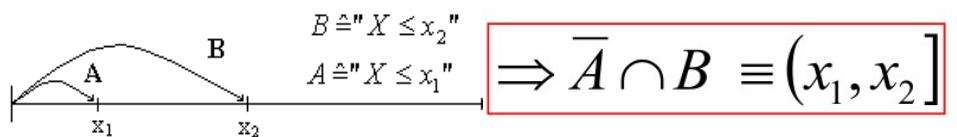
probabilità che la
 variabile aleatoria
 X ha realizzazione
 comprese tra: $x, x + \Delta x$



Punti di discontinuità nella F_X

Si farà riferimento a due risultati, " $X \leq x_1$ " e " $X \leq x_2$ " con $x_1 \leq x_2$, di un esperimento aleatorio inteso come l'osservazione della durata di un certo fenomeno d'attesa, per illustrare il significato e l'utilità della presenza di punti di discontinuità di (salto) prima specie nella classe delle funzioni reali usate per modellare distribuzioni (cumulative) di probabilità nel continuo.

A tal fine, è conveniente far corrispondere gli eventi A e B, rispettivamente, ai due suddetti risultati di durata nel continuo e ragionare sulla seguente figura:



Partendo dalle seguenti relazioni:

$$B = A \cup (\bar{A} \cap B) \Rightarrow P(B) = P(A) + P(\bar{A} \cap B)$$

si riconosce che:

$$P(\bar{A} \cap B) \equiv P(x_1 < X \leq x_2)$$

mentre

$$P(B) - P(A) = P(X \leq x_2) - P(X \leq x_1) = F_X(x_2) - F_X(x_1)$$

quindi

$$P(x_1 \leq X \leq x_2) = F_X(x_2) - F_X(x_1)$$

A questo punto, il fatto che $x_2 \rightarrow x_1$ può essere visto in termini dell'evento $(\bar{A} \cap B)$ - corrispondente ad un intervallo - che tende ad un evento limite corrispondente ad un solo punto della semiretta reale, (Ω è lo spazio dei reali non negativi).

Dunque, l'ipotesi di continuità della F_X nel punto x_1 implica che è nulla la probabilità associata a quel punto (nota anche come *evento limite*) e, viceversa, rinunciando all'ipotesi di continuità della F_X in x_1 , il salto di discontinuità corrisponderebbe alla probabilità finita e non nulla che si vuole associare al punto x_1 .

Per cogliere l'importanza pratica della questione discussa, si può pensare all'esigenza di valutare la probabilità finita di un tempo d'attesa nullo: in tal caso la funzione di distribuzione è continua a destra dello zero.

Prima distribuzione e la più semplice

La distribuzione esponenziale

E' assai importante nell'analisi probabilistica perché **della durata "completamente casuale"** di un fenomeno.

Prima di ricavare l'espressione formale della distribuzione esponenziale, occorre premettere che per "durata completamente casuale" s'intende quella rappresentata da una variabile aleatoria, X, per la quale valga la seguente:

$$Pr\{X \leq t + \Delta t | X \geq t\} = Pr\{X \leq \Delta t\}$$

probabilità che l'evento finisce prima di $t + \Delta t$
sapendo che arriva a t sicuramente è pari
alla probabilità che finisce prima di Δt

Proprietà di assenza di memoria perché NON DIPENDE dalla sua storia

sia $X =$ vita del fenomeno, $Y = X - t$ durata residua del fenomeno e $t =$ "età" del fenomeno

$$Pr\{Y \leq \Delta t | X \geq t\} = Pr\{X \leq \Delta t\}$$

cioè che **la durata residua del fenomeno sia indipendente dall'età, ma distribuita identicamente alla durata vera e propria**. Per un essere umano questo purtroppo non è vero.

Dimostrazione del come arrivare alla distribuzione esponenziale. Ipotesi assenza di memoria Partendo dalla ben nota formula della probabilità congiunta ed applicando la condizionata

$$Pr\{t \leq X \leq t + \Delta t\} = Pr\{X \leq t + \Delta t | X \geq t\} Pr\{X \geq t\}$$

il complemento al minore
uguale è maggiore uguale
per questo ho uguale
imponendo la (*)

$$= Pr\{X \leq \Delta t\} Pr\{X \geq t\}$$

Parto dal rapporto incrementale
per arrivare all'equazione differenziale

dunque:

$$\frac{F(t + \Delta t) - F(t)}{\Delta t} = \frac{F(\Delta t)}{\Delta t} [1 - F(t)] = \frac{F(0 + \Delta t) - F(0)}{\Delta t} [1 - F(t)]$$

$F(0) = 0$ data la proprietà vista
in precedenza

e passando a $\lim_{\Delta t \rightarrow 0}$ si ottiene una semplice equazione differenziale:

$$\frac{d}{dt} F(t) = \frac{d}{dt} F(t)|_{t=0} [1 - F(t)]$$
 è un'equazione differenziale

che ammette come soluzione appunto la legge esponenziale:

$$F(t) = 1 - e^{-\lambda t}, \quad \text{dove} \quad \lambda = \frac{d}{dt} F(t)|_{t=0}$$

lambda è la derivata nell'origine
ovvero che la probabilità finisce subito

Si può vedere l'assenza di memoria come assenza di usura (fattore endogeno), ed il fatto che l'esponenziale termini per qualche fattore esogeno

Un esempio di un fenomeno che non invecchia (privo di memoria) è una richiesta di un servizio.

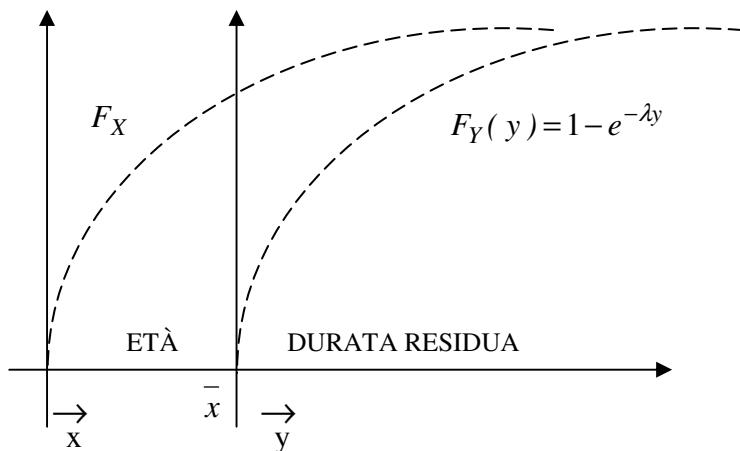
Il ragionamento appena sviluppato porta a definire la distribuzione esponenziale:

$$F_X(x) \hat{=} \begin{cases} 1 - e^{-\lambda x} & x \geq 0 \\ 0 & \text{altrimenti} \end{cases} \quad \text{Non è simmetrica}$$

con la seguente densità esponenziale:

$$f_X(x) = \begin{cases} \lambda e^{-\lambda x} & x \geq 0 \\ 0 & \text{altrimenti} \end{cases}$$

e si dice che per essa vale la proprietà di "assenza di memoria" illustrata dalla seguente figura:



L'esponenziale è l'UNICO che gode di assenza di memoria, che può essere visto come trade-off

DIMOSTRAZIONE DELLA PROPRIETÀ DI ASSENZA DI MEMORIA

Ipotesi: X ha una distribuzione esponenziale

Tesi: $Pr\{Y \leq y | X \geq \bar{x}\} = Pr\{X \leq y\}$ dove $Y = X - \bar{x}$

Prova:

applicando congiunta e condizionata

$$\begin{aligned} P(Y \leq y | X \geq \bar{x}) &= P(X - \bar{x} \leq y | X \geq \bar{x}) = \frac{P(\bar{x} \leq X \leq \bar{x} + y)}{P(X \geq \bar{x})} = \frac{F(\bar{x} + y) - F(\bar{x})}{1 - F(\bar{x})} = \\ &= \frac{1 - \exp\{-\lambda(\bar{x} + y)\} - (1 - \exp\{-\lambda\bar{x}\})}{\exp\{-\lambda\bar{x}\}} \\ &= 1 - \exp\{-\lambda y\} = F_X(y) \hat{=} P(X \leq y) \end{aligned}$$

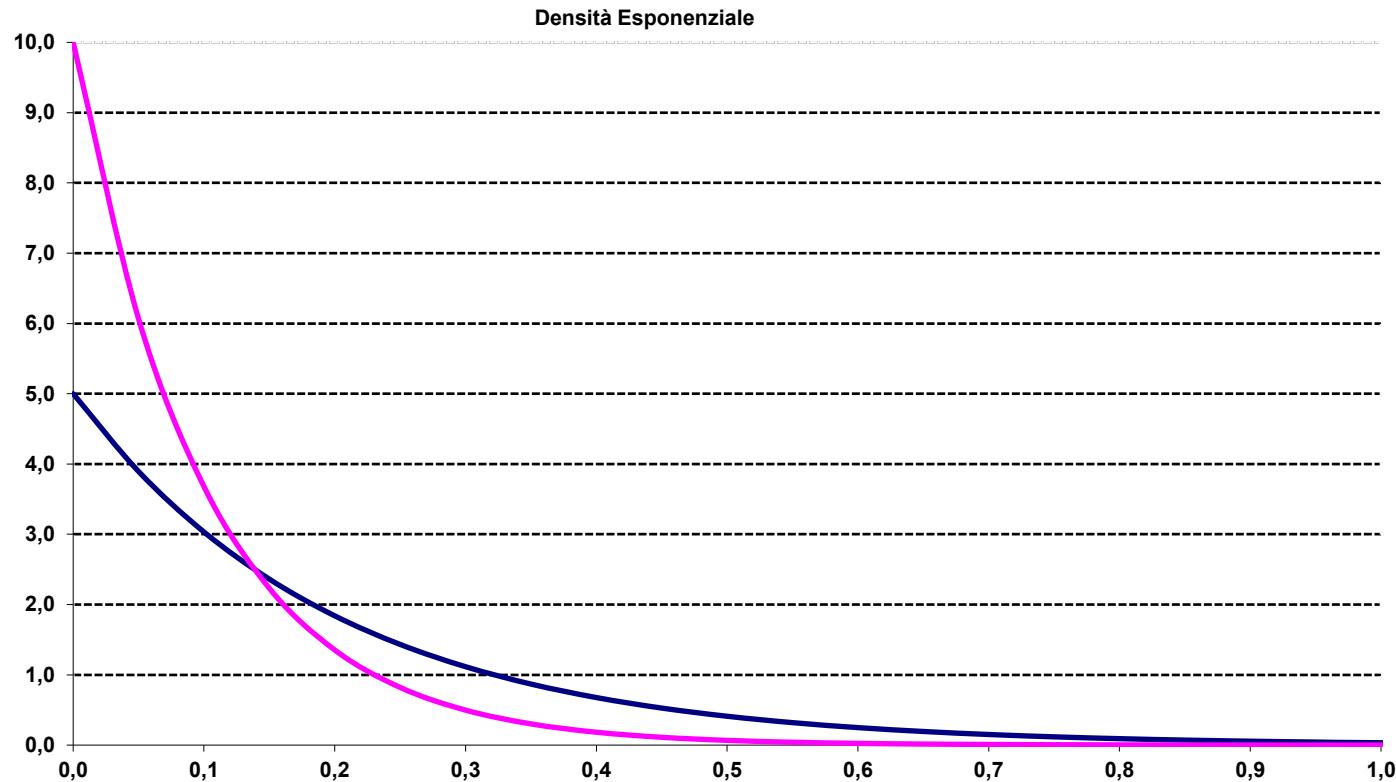
La funzione di distribuzione della vita residua coincide con quella della vita intera, però traslata di una quantità pari all'età.

La funzione densità esponenziale

la blu dura un po' di più della viola
ma comunque tende a finire

$$f_X(x) = \begin{cases} \lambda e^{-\lambda x} & x \geq 0 \\ 0 & \text{altrimenti} \end{cases}$$

esponente negativo
altrimenti sarebbe
crescente all'infinito
e non tendente a 0



l'area sottesa dalla
curva è la probabilità

Esempio

Il funzionamento X di un componente elettronico, inteso come la sua durata, è descrivibile mediante una legge esponenziale $F_X(t) = 1 - e^{-\lambda t}$ dove $\lambda = 0,75$. Posto che il componente sia risultato funzionante all'istante $t = 1$, calcolare la probabilità che lo stesso sia ancora funzionante all'istante $t = 3$.

Partendo dalla formula della probabilità condizionata:

$$\begin{aligned} P(X \leq 3 | X > 1) &= \frac{P(1 < X \leq 3)}{P(X > 1)} = \frac{P(X \leq 3) - P(X \leq 1)}{1 - P(X \leq 1)} = \frac{F_X(3) - F_X(1)}{1 - F_X(1)} \\ &= \frac{1 - e^{-0,75 \cdot 3} - 1 + e^{-0,75 \cdot 1}}{1 - 1 + e^{-0,75 \cdot 1}} \cong \frac{0,895 - 0,528}{0,472} \\ &\cong 0,78. \end{aligned}$$

È possibile pervenire allo stesso risultato sfruttando la proprietà di assenza della memoria della funzione di distribuzione esponenziale:

$$\begin{aligned} P(X \leq 3 | X > 1) &= P(X \leq 2) \\ &= F_X(2) = 1 - e^{-0,75 \cdot 2} \cong 1 - 0,22 \\ &= 0,78. \end{aligned}$$

Affidabilità come misura di durata o di sopravvivenza

Un'immediata applicazione del concetto di funzione di distribuzione nell'analisi probabilistica di sistemi è quella che conduce alla definizione della funzione Affidabilità come misura cumulativa riferita all'intervallo $[0, t]$. (il modello esponenziale si utilizza per l'affidabilità)

A partire dalla variabile aleatoria $X = "tempo al guasto"$, si può definire la funzione affidabilità, $R(t)$, di un componente (o di un sistema):

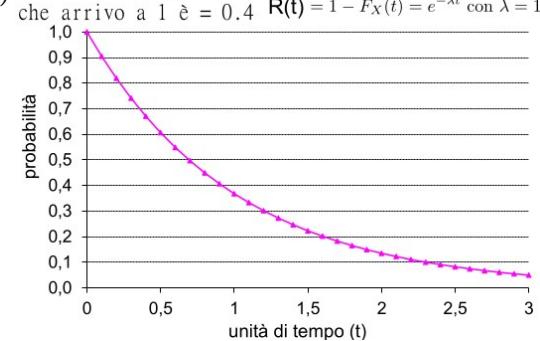
$$R(t) \doteq P(X \geq t) = 1 - P(X \leq t) = 1 - F_X(t) \quad \begin{array}{l} F(t) \text{ è una cumulativa, dunque } R(t) \text{ è} \\ \text{una cumulativa} \\ \Leftarrow \text{perché continua} \end{array}$$

come la probabilità che il componente/sistema abbia una durata di funzionamento che copra l'intervallo $[0, t]$ senza interruzioni, posto che l'istante "0" è l'istante di attivazione del funzionamento. (vuol dire componente collaudato) Ad esempio la probabilità che arrivo a 1 è = 0,4 $R(t) = 1 - F_X(t) = e^{-\lambda t}$ con $\lambda = 1$

Per ipotesi, dunque, deve risultare:

$$R(t \rightarrow 0) \rightarrow 1$$

Proprietà caratteristica è invece:



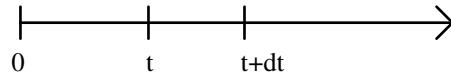
Probabilità del guasto

Faccio la derivata di $R(t)$ così vedo cosa accade in un punto, dunque da

$$\frac{d}{dt} R(t) = \frac{d}{dt} (1 - F_X(t)) = -f_X(t) \quad \begin{array}{l} (\text{essendo l'affidabilità decrescente}) \\ (\text{mi aspetto il meno nel}) \\ (\text{risultato della derivata ricordalo}) \end{array}$$

ottengo che il differenziale di $R(t)$ $f_X(t)dt = -dR(t)$

Tale risultato $-dR(t) = f_X(t)dt$ è noto come "probabilità di un guasto in ($t, t+dt$)".



Più interessante è avere una misura della:

"probabilità di guasto in ($t, t+dt$) | funzionamento fino a t "

che si ottiene introducendo il concetto di tasso di guasto.

A differenza della funzione affidabilità, il tasso di guasto è una misura istantanea e non di tutto l'intervallo precedente il punto considerato. Precisamente, vuole quantificare la

tendenza al guasto che è possibile associare ad un determinato componente/sistema

funzionante in quello stesso istante. (Probabilità che al prossimo istante dopo t si guasti)

Al variare di "t", è rappresentato dalla funzione $h(t)$, definita nel modo seguente:

Dunque, in $h(t)$ per ipotesi si

raggiungerà sicuramente

l'istante t , in particolare $h(t) \triangleq \lim_{\Delta t \rightarrow 0} \frac{\text{Prob}(t < X \leq t + \Delta t | X > t)}{\Delta t}$

ignorando delta t
e sostituendo con
congiunta e condizionante

che si raggiunge si chiama età

$$\text{risulta che } h(t) = \lim_{\Delta t \rightarrow 0} \frac{\text{Prob}(t < X < t + \Delta t)}{\Delta t P(X > t)} = \lim_{\Delta t \rightarrow 0} \frac{F_X(t + \Delta t) - F_X(t)}{\Delta t R(t)}$$

rapporto incrementale

$$\text{ovvero } f_X(t) = \lim_{\Delta t \rightarrow 0} \frac{F_X(t + \Delta t) - F_X(t)}{\Delta t (1 - F_X(t))} = \frac{f_X(t)}{1 - F_X(t)} = \frac{\lambda e^{-\lambda t}}{1 - (1 - e^{-\lambda t})} = \lambda$$

λ del esponenziale, è il rischio del guasto ovvero il parametro che dice quanto velocemente finisce una cosa. Sostituendo correttamente $1 - F_X(t) = R(t)$

si ottiene che $h(t) = \lim_{\Delta t \rightarrow 0} \frac{R(t) - R(t + \Delta t)}{\Delta t R(t)} = \frac{-R'(t)}{R(t)}$ che messa assieme a $f_X(t)dt = -dR(t)$
h(t) sarebbe una velocità. porta a $R(t) \cdot h(t)dt = f_x(t)dt$
h(t) dt è probabilità

Da quest'ultima è possibile ottenere una relazione che lega il tasso stesso all'affidabilità, integrando il primo e il secondo membro tra 0 e u :

$$\int_0^u h(t)dt = \int_0^u -\frac{R'(t)}{R(t)}dt = -\int_{R(0)}^{R(u)} \frac{dR(t)}{R(t)}$$

ovvero, poiché $R(0) = 1$,

$$-\ln R(u) = \int_{t=0}^u h(t)dt,$$

ed in conclusione

funzione di rischio cumulato

$$\Rightarrow R(u) = \exp \left[- \int_{t=0}^u h(t)dt \right] \Rightarrow 1 - F_X(u) = \exp \left[- \int_0^u h(t)dt \right]$$

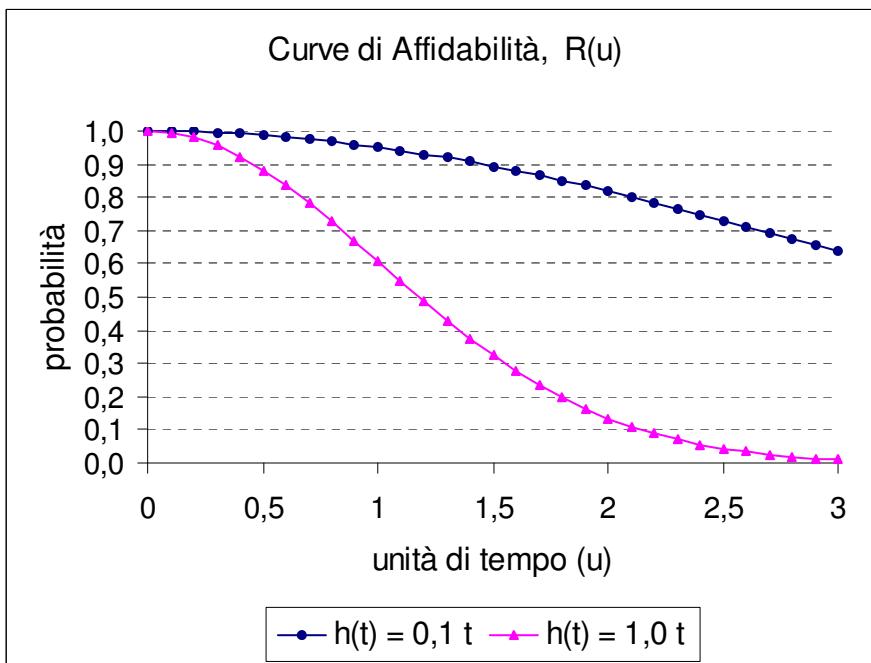
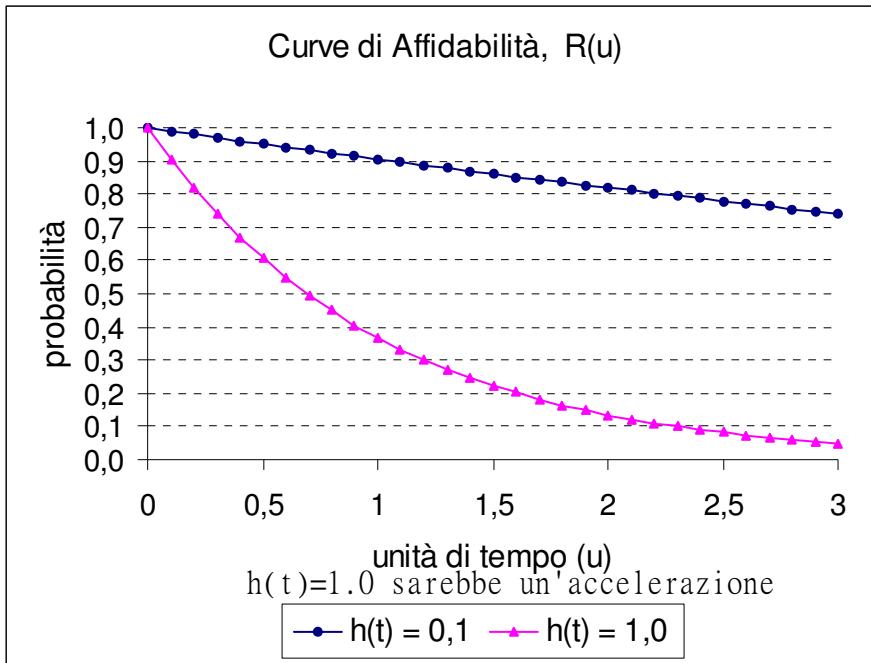
Caso particolare: $h(t) = \lambda \Rightarrow R(u) = \exp\{-\lambda \cdot u\}$

La funzione $h(t)$ è il "modello di guasto" alla base della funzione affidabilità.

(se mi dai la velocità con cui si consuma un fenomeno io so dove ti trovi)

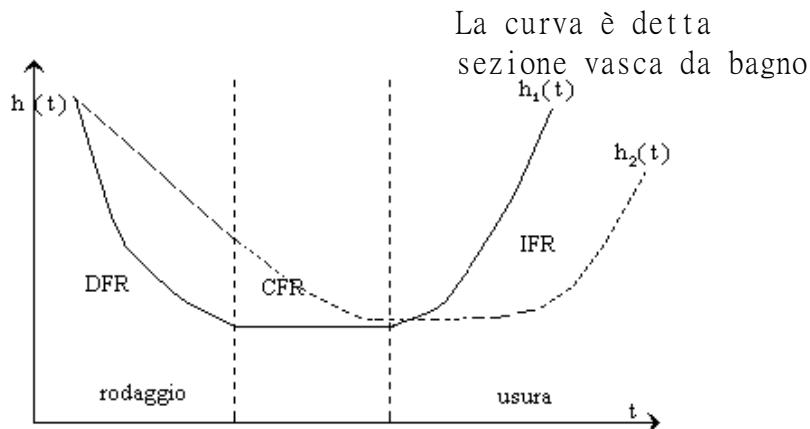
Il rischio di guasto può quindi essere o meno costante. Se è costante l'affidabilità si modella come una semplice esponenziale altrimenti al posto dell'esponente si trova l'integrale della funzione di rischio

Tasso di guasto costante e tasso lineare



Modelli di guasto

DFR=decreasing failure rate
 CFR=constant failure rate
 IFR= increasing failure rate



(DECREASING-CONSTANT-INCREASING FAILURE RATE)

Il grafico raffigura andamenti verosimili che possono essere attribuiti alla funzione $h(t)$

$h_1(t)$: limitato periodo di rodaggio seguito da un periodo a tasso di guasto costante, detto vita utile, e poi dal periodo di usura;
 $h_2(t)$: periodo di rodaggio più lungo e poi vita utile e usura.

Weibull è un'estensione del caso precedente particolare $h(t) = \lambda$

Il modello in qualche maniera rappresenta la curva del grafico

Il modello di Weibull: $h(t) \doteq \lambda \alpha t^{\alpha-1}$, $\lambda > 0, \alpha > 0$.

Proprietà: $(\alpha > 1 \Rightarrow IFR, \alpha < 1 \Rightarrow DFR, \alpha = 1 \Rightarrow CFR)$

Grazie alla relazione:

$$R(u) = \exp \left[- \int_{t=0}^u h(t) dt \right] = 1 - F_X(u)$$

Porta alla definizione delle funzioni densità e distribuzione di Weibull:

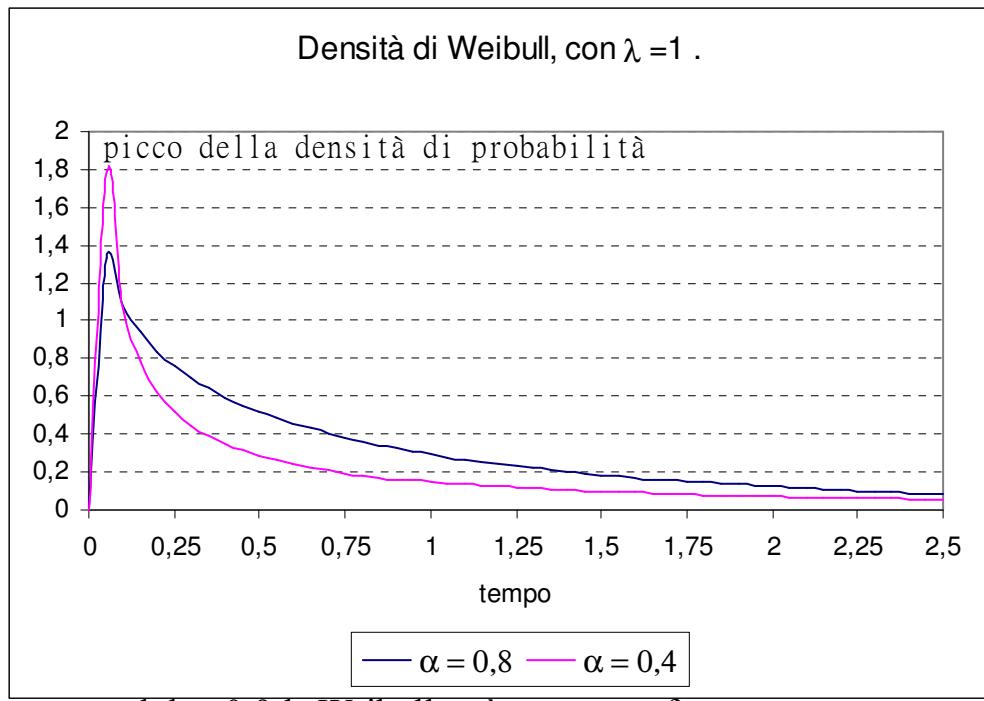
$$f_X(t) \doteq \lambda \alpha t^{\alpha-1} \exp \left\{ - \lambda t^\alpha \right\}, \quad \lambda > 0, \alpha > 0, \quad t \geq 0$$

$$F_X(t) \doteq 1 - \exp \left\{ - \lambda t^\alpha \right\}, \quad \lambda > 0, \alpha > 0, \quad t \geq 0$$

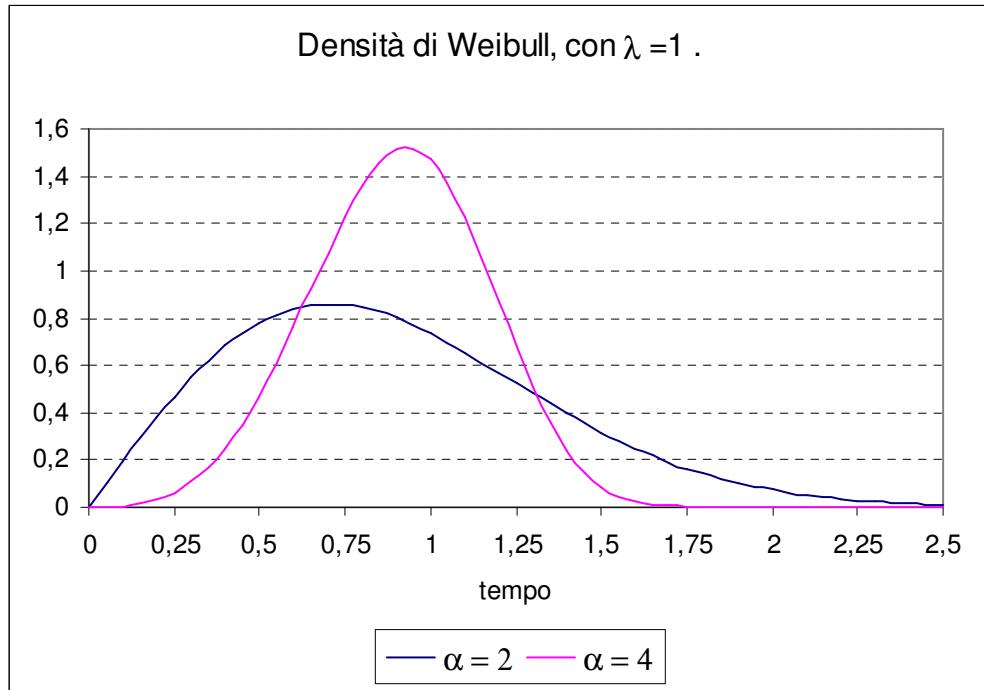
Weibull può descrivere la durata di vita per un fenomeno la cui "probabilità di morire" può variare nel tempo, a differenza della distribuzione esponenziale che prevede tassi di guasto costanti nel tempo.

Rappresentazione grafica del modello di Weibull

utilizzato nella mortalità infantile



con alpha=0.9 la Weibull può essere confusa con un esponenziale



con $\alpha > 1$ si creano
curve totalmente diverse

Gli istogrammi con Weibull sono complicati e se $\alpha \leq 1$, risultano ancora più complicati

Media, varianza e momenti di una variabile aleatoria

Con il concetto di media di una variabile aleatoria si vuole associare un singolo numero alla "forma" secondo la quale la probabilità è ripartita fra le singole realizzazioni della variabile stessa.

A tal fine, con riferimento ad una variabile aleatoria discreta e non negativa, ad es. N a valori in $[0,1,2,\dots,n,\dots]$ si definisce momento del primo ordine, o valore atteso, o speranza matematica, o, più semplicemente media e si indica con $E[N]$ il valore finito della seguente serie, quando esiste:

$$\sum_{n=0}^{\infty} n P_N(n) \triangleq E[N]$$

Una media aritmetica stima un valore atteso ecco perchè sono sinonimi

Con riferimento a X non negativa e continua a valori in $[0, \infty)$, si definisce alla stessa maniera e si indica con $E[X]$ il valore numerico finito del seguente integrale, quando esiste:

L'integrale è la somma pesata del prodotto tra realizzazione e probabilità che si verifichi

$$\int_{x=0}^{\infty} x f_X(x) dx \triangleq E[X]$$

$E[X]$ è un numero!
Proprietà:

$E[\text{numero}] = \text{numero}$ quindi
 $E[E[X]] = \text{numero!}$

Generalizzando, sono detti momenti di ordine "k" della X e della N, rispettivamente, i seguenti:

$$E[X^k] \triangleq \int_{x=0}^{\infty} x^k f_X(x) dx, \quad E[N^k] \triangleq \sum_{n=0}^{\infty} n^k P_N(n)$$

Con il concetto di varianza, invece, si vuole introdurre una misura numerica dell'entità della dispersione delle realizzazioni possibili attorno alla media.

Dunque si definisce: è lo scostamento dal valore atteso: appunto valore atteso x meno qualcosa

$$VAR[X] \triangleq E[(x - E[X])^2] = \int_{x=0}^{\infty} (x - E[X])^2 f_X(x) dx$$

x-E[X] è la variabile aleatoria detta scostamento

non si preoccupa dove si discostano gli scarti della differenza

nel caso discreto $VAR[N] \triangleq E[(n - E[N])^2] = \sum_{n=0}^{\infty} (n - E[N])^2 P_N(n)$

P.S. calcolare il valore atteso di Weibull è complicato dato che devi fare un integrale (esula dal corso)

La varianza di una somma di variabili aleatorie è pari alla somma delle varianze di ciascuna se assumiamo che quelle variabili siano indipendenti mentre per il valore atteso di una somma di variabili aleatoria non deve valere l'indipendenza.

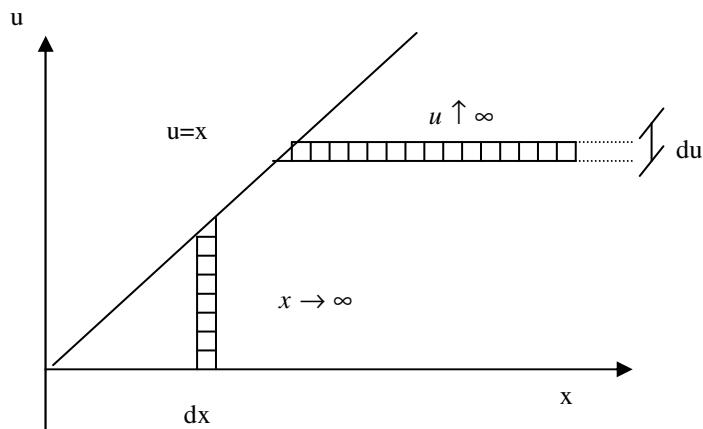
Formule alternative per il calcolo del valore atteso

Per il calcolo del valore atteso di una variabile aleatoria non negativa, vale la seguente formula alternativa:

$$E[X] \hat{=} \int_{x=0}^{\infty} [1 - F_X(x)] dx \quad \text{utilizzando l'affidabilità } R(t)$$

PROVA:

$$\begin{aligned} \int_{x=0}^{\infty} xf_X(x) dx &= \int_{x=0}^{\infty} \int_{u=0}^x du f_X(x) dx = \\ &= \int_{u=0}^{\infty} \int_{x=u}^{\infty} f_X(x) dx du = \int_{u=0}^{\infty} [1 - F_X(u)] du \end{aligned}$$



Per il calcolo del valore atteso della N, vale la seguente formula alternativa:

$$E[N] \hat{=} \sum_{n=0}^{\infty} n P(N=n) = \sum_{n \geq 1} P(N \geq n)$$

Infatti:

$$\begin{aligned} \sum_{n=0}^{\infty} n P_N(n) &= \\ P_N(1) + P_N(2) + P_N(2) + P_N(3) + P_N(3) + P_N(3) + \dots &= \sum_{n \geq 1} P(N \geq n) \end{aligned}$$

Calcolo del valore atteso della legge esponenziale

Si possono seguire due metodi alternativi:

$$A) \quad E[X] \hat{=} \int_{t=0}^{\infty} t \lambda e^{-\lambda t} dt$$

ovvero risolvendo per parti

$$\frac{1}{\lambda} \int t \lambda e^{-\lambda t} d(\lambda t) = \frac{1}{\lambda} \int x e^{-x} dx \quad (\text{avendo posto: } x \hat{=} \lambda t)$$

Poiché:

$$d(xe^{-x}) = xd(e^{-x}) + e^{-x}dx = -xe^{-x}dx + e^{-x}dx$$

$$\int d(xe^{-x}) = - \int xe^{-x}dx + \int e^{-x}dx$$

$$\Rightarrow \int xe^{-x}dx = - \int d(xe^{-x}) + \int e^{-x}dx =$$

$$[-xe^{-x}]_0^\infty + [-e^{-x}]_0^\infty = 0 + 1$$

si ottiene:

$$\frac{1}{\lambda} \int_0^\infty xe^{-x}dx = \frac{1}{\lambda} = E[X]$$

IN PARTICOLARE $\frac{1}{\lambda}$ È DETTO MTTF (MEAN TIME TO FAILURE) TEMPO MEDIO AL GUASTO

OSSERVAZIONE:

Il reciproco del valore atteso della variabile aleatoria X è quell'unico parametro, λ , che compare nella definizione della distribuzione esponenziale $F_X(x) = 1 - \exp\{-\lambda x\}$ e nell'espressione della densità $f_X(x) = \lambda \exp\{-\lambda x\}$.

Calcolo del momento del secondo ordine e della varianza

È stato appena dimostrato che la media di una variabile aleatoria distribuita con legge esponenziale di parametro λ è pari a $1/\lambda$. Dunque, è bene puntualizzare che se X rappresenta il tempo di vita di un componente non soggetto ad usura, allora il tempo medio al guasto è:

$$E[X] = \int_0^\infty R(t)dt = \frac{1}{\lambda}, \quad \text{con } R(t) = \exp\{-\lambda t\}$$

e, nella notazione anglosassone dell'affidabilità, è noto con l'acronimo **MTTF** (mean time to failure).

Dalla definizione:

$$E[X^2] = \int_0^\infty t^2 f_X(t)dt$$

ricordando che:

$$f_X(t)dt = dR(t) = R'(t)dt$$

si ottiene:

$$\begin{aligned} E[X^2] &= - \int_0^\infty t^2 R'(t)dt \\ &= -t^2 R(t) \Big|_0^\infty + \int_0^\infty 2t^{2-1} R(t)dt = \int_0^\infty 2t R(t)dt \end{aligned}$$

e di nuovo, con $R(t) = \exp\{-\lambda t\}$, risulta: $E[X^2] = \frac{2}{\lambda^2}$.

Per calcolare la varianza (del tempo di vita) si può usare la formula seguente, che sarà dimostrata più avanti:

$$\begin{aligned} VAR[X] &= E[X^2] - E[X]^2 && \text{(Utile in futuro)} \quad \text{(varianza è momento del secondo ordine meno momento del primo ordine al quadrato)} \\ &= \int_0^\infty 2t R(t)dt - \left(\int_0^\infty R(t)dt \right)^2 \end{aligned}$$

e per la legge esponenziale risulta: $VAR[X] = \frac{2}{\lambda^2} - \frac{1}{\lambda^2} = \frac{1}{\lambda^2}$

La varianza dell'esponenziale è molto grande

EXTRA - MEDIA E VARIANZA DI WEIBULL

ALPHA E BETA SONO DETTI RISPETTIVAMENTE
PARAMETRI DI FORMA E DI SCALA

Proposizione :

Sia X una var. al. di Weibull con parametri α e β ($\beta \triangleq \frac{1}{\Gamma(1/\alpha)}$) secondo notazione di Excel.

Allora risultano:

$$\text{valore atteso} \triangleq \mu = \bar{\alpha}^{1/\beta} \Gamma(1+1/\beta)$$

$$\text{varianza} \triangleq \sigma^2 = \bar{\alpha}^{-2/\beta} \Gamma(1+2/\beta) - \mu^2$$

$$(\text{ricordare: } \sigma^2 = E[X^2] - E[X]^2)$$

Prova : Partiamo dalla espressione della densità di Weibull usata da EXCEL

$$f_X(x) \triangleq \alpha \cdot \beta \cdot x^{\beta-1} \cdot \exp(-\alpha x^\beta)$$

Calcoliamo il valore atteso:

$$E[X] = \int_{x=0}^{\infty} x \cdot \alpha \beta x^{\beta-1} \cdot e^{-\alpha x^\beta} \cdot dx$$

$$= \int_{x=0}^{\infty} \alpha \beta x^{\beta} \cdot e^{-\alpha x^\beta} \cdot dx$$

$$\text{Poniamo } z = \alpha x^\beta$$

$$\text{quindi} \Rightarrow x = (z/\alpha)^{1/\beta} \quad dx = \frac{1}{\alpha \beta} \left(\frac{z}{\alpha}\right)^{\frac{1}{\beta}-1} dz$$

$$\begin{aligned}
 \text{Da qui } \Rightarrow E[X] &= \int_{z=0}^{\infty} \alpha \beta \frac{z}{\alpha} \cdot e^{-z} \cdot \frac{1}{\alpha \beta} \cdot \left(\frac{z}{\alpha}\right)^{\frac{1}{\beta}-1} \cdot dz \\
 &= \int_0^{\infty} \left(\frac{z}{\alpha}\right)^{\frac{1}{\beta}} \cdot e^{-z} \cdot dz \\
 &= \alpha^{-\frac{1}{\beta}} \int_{z=0}^{\infty} z^{\frac{1}{\beta}} e^{-z} \cdot dz
 \end{aligned}$$

Ebbene, l'integrale indefinito che vedete è la funzione "gamma" di Euler (precisamente: il II integrale di Euler) che si riporta sui libri con la seguente notazione:

$$M(\alpha) \stackrel{\Delta}{=} \int_{z=0}^{\infty} z^{\alpha-1} \cdot e^{-z} \cdot dz \quad \text{con } \alpha > 0$$

Il calcolo di $M(\alpha)$ per specifici valori di α è detto "quozio di Euler" ☺

Nel nostro caso, ci limitiamo ad osservare quanto segue:

$$\begin{aligned}
 \int_{z=0}^{\infty} z^{\frac{1}{\beta}} \cdot e^{-z} \cdot dz &\equiv M\left(1 + \frac{1}{\beta}\right) \Rightarrow \text{conclusione:} \\
 &\Rightarrow E[X] = \alpha^{-\frac{1}{\beta}} \cdot M\left(1 + \frac{1}{\beta}\right)
 \end{aligned}$$

Statistiche dell'Ordinamento

Date X_1, X_2, \dots, X_n variabili aleatorie indipendenti ed identicamente distribuite, sono utili

nell'analisi probabilistica le variabili aleatorie Y_1, \dots, Y_n definite alla seguente maniera:

Presi n studenti che fanno uno scritto che partono allo stesso momento $Y_1 \hat{=} \min\{X_1, X_2, \dots, X_n\}$ Y_1 è la statistica del primo ordine e corrisponde al primo studente che consegna

e così via $Y_2 \hat{=} \min[\{X_1, X_2, \dots, X_n\} - \{Y_1\}] \dots Y_{n-1} \hat{=} \min[\{X_1, X_2, \dots, X_n\} - \{Y_1, Y_2, \dots, Y_{n-2}\}]$

Dunque Y_n corrisponde $Y_n = \max\{X_1, X_2, \dots, X_n\}$ sarebbero durate che partono all'ultimo che consegna tutte allo stesso momento perciò

Y_i sono funzioni di note come statistiche dell'ordinamento. Infatti, detta Y_k la statistica di ordine k di variabili aleatorie

($k=1, \dots, n$), è immediato riconoscere che, se X_i rappresenta il tempo al guasto del componente i ($i=1, \dots, n$), allora la statistica di ordine $k = n-m+1$ rappresenta il tempo al guasto di un sistema "m out of n". $\min/\max\{X_1, X_2, \dots, X_n\}$ è una funzione di variabili aleatorie come si vedrà

In linea di principio, per calcolare l'affidabilità del sistema occorrerebbe determinare la funzione di distribuzione della statistica corrispondente, dato che:

Y_i può essere anche visto come

l'i-esima richiesta ad un server

$$R_{m|n}(t) = 1 - F_{Y_{n-m+1}}(t)$$

In pratica, note le funzioni di affidabilità $R_1(t), \dots, R_n(t)$ di tutti i componenti, è possibile ricavare direttamente l'affidabilità del sistema con le prove di Bernoulli.

Il sistema "m out of n" risulterà funzionante all'istante t se risulterà funzionante *almeno uno* dei campioni di *almeno m* componenti che è possibile estrarre dalla popolazione di numerosità pari a n . Dunque, riconoscendo che:

- $\prod_{i=1}^{n-j} [1 - R_i(t)] \prod_{i=1}^j R_i(t)$ è la probabilità che risultino funzionanti all'istante t un campione di j componenti indipendenti; tra $n-j$ non funzionano e j che funzionano
- $\binom{n}{j}$ è il numero di campioni di numerosità j che devono essere considerati;
- j deve variare da m a n , in modo da considerare i campioni di tutte le dimensioni utili;

si ottiene. Il caso più semplice dove i componenti sono identici:

$$R_{m/n}(t) = \sum_{j=m}^n \binom{n}{j} [1 - R(t)]^{n-j} R(t)^j$$

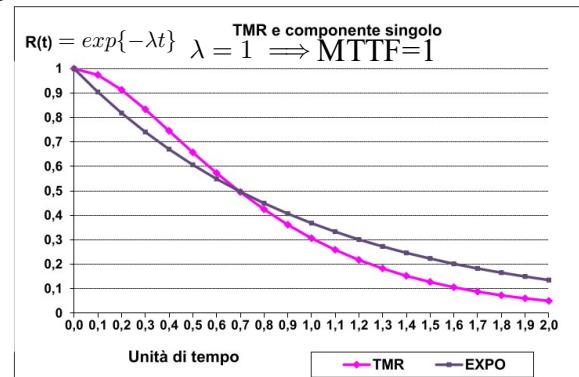
Casi particolari

- Il TMR con componenti identici ($R_1(t) = \dots = R_n(t) \hat{=} R(t)$). "perfetto"

Assumendo che il comparatore abbia probabilità di guasto trascurabile :

$$R_{TMR}(t) = R_{2/3}(t) = \binom{3}{2} [1 - R(t)][R(t)]^2 + \binom{3}{3} [R(t)]^3 = 3[R(t)]^2 + 2[R(t)]^3$$

La probabilità in blu è una congiunta ed è la probabilità che il sistema funziona a con 2 su 3 mentre la verde è che il modello funzioni 3 su 3. Il grafico mostra che la TMR interseca l'esponenziale



Affidabilità sistema parallelo= "1-out-of n":

$$R_{1/n}(t) = \sum_{j=1}^n \binom{n}{j} [1 - R(t)]^{n-j} R(t)^j = 1 - F_{Y_n}(t) \quad \text{Con } Y_n = \max\{X_1, X_2, \dots, X_n\} \quad \text{procedendo con la dimostrazione si ha che}$$

$$R_{1/n}(t) = \sum_{j=1}^n \binom{n}{j} [1 - R(t)]^{n-j} R(t)^j = \sum_{j=0}^n \binom{n}{j} [1 - R(t)]^{n-j} R(t)^j - \binom{n}{0} [1 - R(t)]^n = [1 - R(t) + R(t)]^n - 1 \cdot [1 - R(t)]^n = 1 - [1 - R(t)]^n \quad \text{esplicitando le mortalità: } [1 - R(t)]^n \implies 1 - Pr\{X_1 \leq t \cap \dots \cap X_n \leq t\} = 1 - \prod_{i=1}^n F_{X_i}(t) \quad \text{Eguagliando i due risultati arancioni } F_{Y_n}(t) = \prod_{i=1}^n F_{X_i}(t) \quad \text{perciò la funzione tempo al guasto del ultimo è dato dal prodotto delle distribuzioni.}$$

Affidabilità sistema seriale= "n-out-of n":

$$R_{n/n}(t) = Pr\{X_1 > t \dots X_n > t\} = \prod_{i=1}^n R(t) = \binom{n}{n} [1 - R(t)]^{n-n} R(t)^n = 1 - F_{Y_1}(t) \quad \text{con } Y_1 = \min\{X_1, X_2, \dots, X_n\}$$

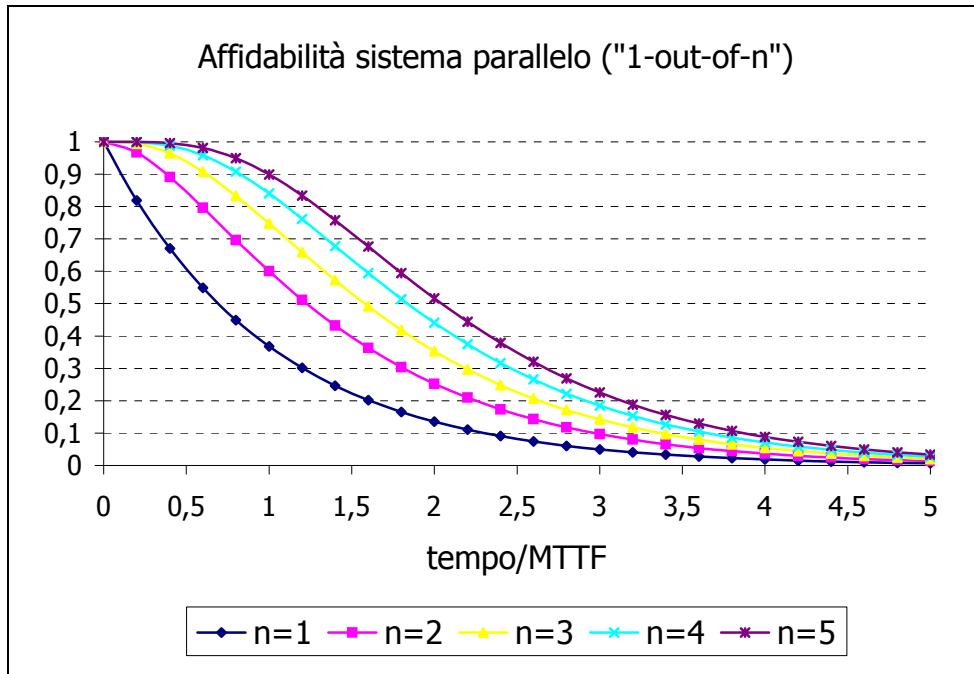
Eguagliando i risultati in fucsia, ed esplicitando R(t) si ottiene che: $\prod_{i=1}^n [1 - F_{X_i}(t)] = 1 - F_{Y_1}(t)$ ovvero che la produttoria fra le complementari è pari alla sopravvivenza.

Come corollario al calcolo dell'affidabilità del sistema seriale, si può ricavare un risultato utile in generale:
una funzione di distribuzione definita come min di altre n esponenziali indipendenti è anch'essa esponenziale, con parametro pari alla somma degli n parametri. Infatti, se:

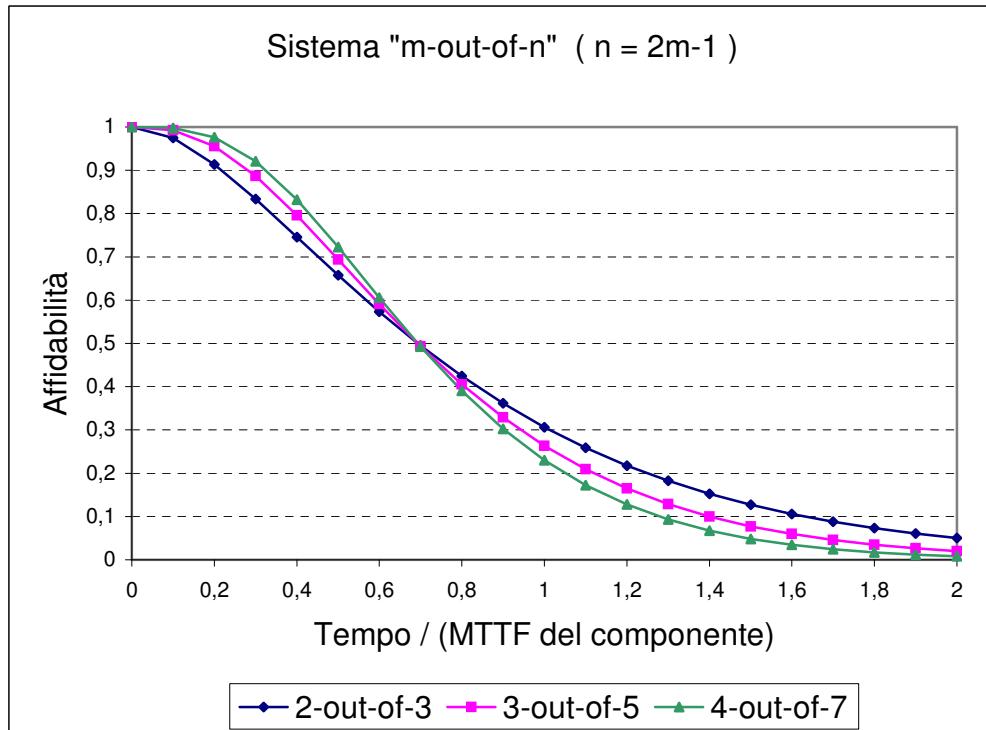
$$F_{X_i}(t) = 1 - e^{-\lambda_i t} \implies F_{Y_1}(t) = 1 - \prod_{i=1}^n e^{-\lambda_i t} = 1 - \exp[-\sum_{i=1}^n \lambda_i]t$$

Il min rappresenta nell'affidabilità il tempo minimo per un guasto. Nel caso di componenti in serie il min rappresenta il tempo al guasto del sistema!

Grafici di affidabilità sistemi "m-out-of-n"



dopo 4 componenti in parallelo l'affidabilità che si guadagna non è poi così tanta

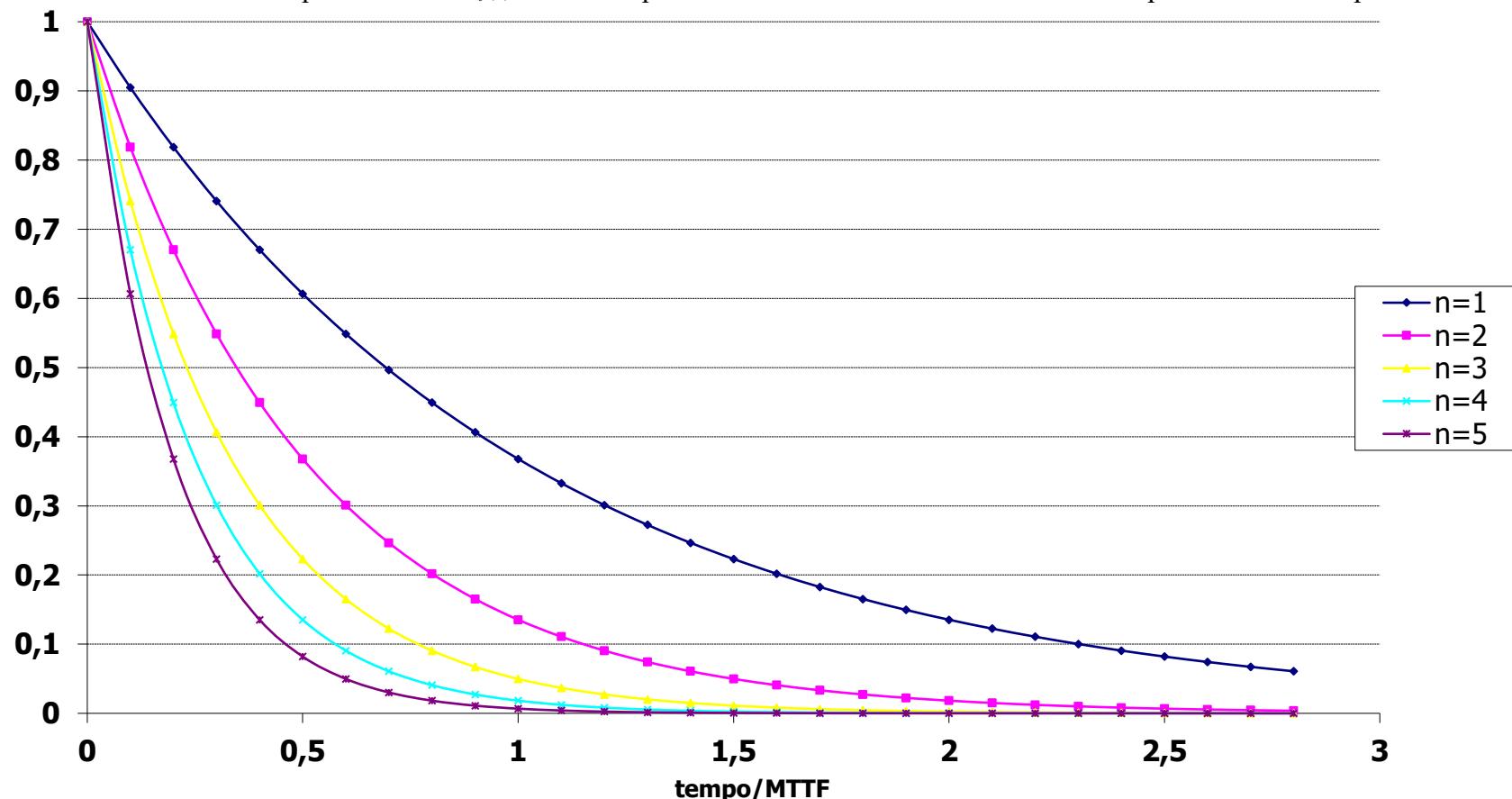


Il 2 su 3 a lungo andare è quello che ha affidabilità più alta anche se all'inizio è quello che perde affidabilità più velocemente.

AFFIDABILITA' DEL SISTEMA SERIALE A COMPONENTI IDENTICI

Affidabilità sistema seriale ("n-out-of-n")

In questo caso $F_{X_i}(t) = 1 - e^{-\lambda t}$ quindi mostra come il modello seriale con l'esponenziale si va a perdere subito



Esercizio

Sistema “m-out-of-n”

- L'elaborazione di alcuni dati acquisiti da un aeromobile durante il volo viene fatta da quattro processori indipendenti che lavorano in parallelo.
- Ad ognuno viene dato lo stesso input e si suppone che i risultati siano corretti se almeno tre dei quattro processori restituiscono lo stesso risultato.
- Il comparatore dei risultati può essere considerato "perfetto" e la legge di guasto, identica per i quattro componenti, è quella esponenziale, con un MTTF pari a 10.000 ore:

Esercizio – Quesiti (1_di_2)

1. Calcolare la probabilità che nessuno dei quattro componenti si guasti prima del 5% del proprio MTTF.
Quindi che supera il 5%
2. E la probabilità che il sistema superi le *500 ore di vita?*
(affidabilità)
3. Qual è, invece, la probabilità che *dopo le 500 ore* di vita risultino funzionanti *non più di due* dei quattro componenti?
Quindi al massimo due processori qualunque 0 e 1 sono compresi!
4. E il tasso di guasto del sistema alle *500 ore di vita?*
5. Ricalcolare il tasso di guasto del sistema supponendo, però, che alle 500 ore il sistema si sia appena ridotto a 3 componenti, per effetto di uno ed un solo guasto.

Esercizio

Soluzione del quesito 1 (1_di_2)

Sia X_i il tempo al guasto del componente i-esimo ($i=1,2,3,4$)

Sappiamo che:

$$\lambda_i = \frac{1}{10000} = 10^{-4} \text{ ore}^{-1} \quad \forall i = 1, \dots, 4$$

ossia, $\lambda_i = \lambda$ ($i=1, \dots, 4$).

Pertanto le funzioni di distribuzione saranno:

$$F_{X_i}(t) = 1 - e^{-\lambda t}$$

Esercizio

Soluzione del quesito 1 (2_di_2)

Definita la v.a.

$$X = \min\{X_1, X_2, X_3, X_4\}$$

Possiamo calcolare la probabilità che nessuno dei quattro componenti si guasti prima del 5% del proprio MTTF come:

$$P(X > t) = 1 - F_X(t) = \exp\left[-t \sum_{i=1}^4 \lambda_i\right]$$

Sostituendo i valori noti e ricordando che $\lambda_i = \lambda$ ($i=1, \dots, 4$) si ottiene:

$$P(X > 500) = e^{-4 \cdot 10^{-4} \cdot 500} = 0.819$$

Esercizio

Soluzione del quesito 2

L'affidabilità per $t=500$ ore di un singolo componente è:

$$R(t = 500) = P(X > 500) = e^{-10^{-4} \cdot 500} = 0.951$$

Allora la probabilità che il sistema risulti funzionante alle 500 ore di vita è:

$$\begin{aligned} R_{3/4}(t = 500) &= \sum_{j=3}^4 \binom{4}{j} \left[e^{-10^{-4} \cdot 500} \right]^j \left[1 - e^{-10^{-4} \cdot 500} \right]^{4-j} \\ &= 4 \cdot [R(t = 500)]^3 - 3 \cdot [R(t = 500)]^4 \\ &= 0.986 \end{aligned}$$

Esercizio

Soluzione del quesito 3

L'affidabilità per $t=500$ ore di un singolo componente è:

$$R(t = 500) = P(X > 500) = e^{-10^{-4} \cdot 500} = 0.951$$

Dunque, la probabilità che a 500 ore (da inizio missione) risultino funzionanti **non più** di due dei quattro componenti è la seguente:

$$\begin{aligned} R_{\leq 2/4}(t = 500) &= \sum_{j=0}^2 \binom{2}{j} \cdot \left[e^{-10^{-4} \cdot 500} \right]^j \cdot \left[1 - e^{-10^{-4} \cdot 500} \right]^{2-j} \\ &= 1 \cdot 1 \cdot \left[1 - e^{-10^{-4} \cdot 500} \right]^2 + 2 \cdot \left[e^{-10^{-4} \cdot 500} \right] \cdot \left[1 - e^{-10^{-4} \cdot 500} \right] + \\ &\quad + 1 \cdot \left[e^{-10^{-4} \cdot 500} \right] \cdot 1 = 0.013 \end{aligned}$$

in particolare è pari al complementare del quesito 3 perché $1 - 0.986 = 0.013$. Solo che in questo esercizio vale perché "per caso" il sistema funziona se 3 su 4 funzionano

Esercizio Soluzione del quesito 4

Il tasso di guasto del sistema $h(t)$ alle prime 500 ore di vita si calcola da:

$$h(t) = \frac{f_X(t)}{1 - F_X(t)}$$

ricorda le definizioni di $h(t)$
e di $F_X(t)$

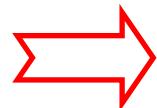
$F_X(t)$ nel caso specifico risulta:

$$F_X(t) = 1 - 4 \exp[-3\lambda t] + 3 \exp[-4\lambda t]$$

Da qui, $f_X(t)$ si ottiene derivando:

$$f_X(t) = 12 \lambda \exp[-3\lambda t] - 12 \lambda \exp[-4\lambda t]$$

Con $h(t)$ sarebbe da fare uno studio di funzione
per poi ottenere



$$h(t=500) = 0,000051$$

Esercizio

Soluzione del quesito 5

Le 500 ore sarebbero l'età e ricordando che la vita residua non dipende dall'età, quindi il rischio di guasto rimane uguale a prima

Nel caso in cui, a $t=500$ ore, si sia appena verificato il primo guasto, il sistema è diventato la serie di tre componenti. Cambia quindi la $F_X(t)$:

$$F_X(t) = 1 - \exp[-3\lambda t]$$

Il tasso di guasto, $h(t)$, si determina ricordando la proprietà di assenza di memoria della distribuzione esponenziale:

Ricordando che il rischio di guasto di UN componente è λ
Allora il rischio del sistema sarà la somma dei 3

$$h(t) = 3\lambda$$

Devi capire bene lambda

Non c'è bisogno della seguente:

$$h(t) = \frac{f_X(t)}{1 - F_X(t)}$$

Esercizio – Quesiti (2_di_2)

- 6) Effettuare lo studio di funzione su $h(t)$ al quesito 4
- 7) Calcolare un limite inferiore all'affidabilità del sistema all'istante $t = 500h$ vedi limiti al inizio
- 8) Definire la var. al. "tempo (residuo) al guasto del sistema" una volta che, per effetto del 1° guasto accaduto a $t=500h$, è stato deciso di riconfigurare il sistema come un "2-out-of-4". che accade?

Soluzione: lasciata allo studente ☺

Esercizio 1

(Problema e quesiti)

Il tasso di guasto dell'esponenziale è costante

Il “*rischio di guasto*” di un componente **sta crescendo** linearmente col tempo d'uso, secondo un coefficiente pari a $0,2 \text{ (u.t.}^{-2}\text{)}$.

da notare elevato a -2 e non -1 il tasso di guasto

1. Proporre un modello per la valutazione della affidabilità del componente. ovviamente esponenziale non va bene perché non è costante il rischio di guasto
2. Calcolare la probabilità che il componente funzioni ancora dopo 2 unità di tempo. (Affidabilità)
3. Calcolare la probabilità di un guasto nell'istante di tempo 0,25 e in un intorno $dt=0,01$ ovvero $[0,25; 0,26]$.

Esercizio 1

(Soluzione dei quesiti 1 e 2)

1. Si adotta un modello di guasto lineare

$$h(t) = 0,2 \cdot t$$

2. L'affidabilità si valuta nel modo seguente:

Applico la relazione tra tasso di guasto ed affidabilità

$$\begin{aligned} R(u) \Big|_{u=2} &= P(X \geq 2) = \exp \left[- \int_{t=0}^2 h(t) dt \right] = \\ &= \exp \left[- \int_{t=0}^2 0.2t dt \right] = \exp(-0.1 \cdot u^2) \Big|_{u=2} = 0.6703 \end{aligned}$$

Esercizio 1

(Soluzione del quesito 3)

Nella probabilità di guasto non si sa se arrivo a t. Mentre per il tasso di guasto è garantito che si arriva a t

3. La probabilità di un guasto in (t) si calcola come:

F è differenziabile per ipotesi

Per la probabilità uso $f(t)dt$
per il tasso uso $h(t)dt$

$$P(X = t) = dF(t) = f(t)dt \quad \text{con } dt = 0 \quad \rightarrow \quad P(X = t) = 0$$

La probabilità di un guasto in $[t; t+dt]$ si calcola come:

$$P(t \leq X \leq t + dt) = f(t) \cdot dt \quad \text{ricordando la relazione tra } f(t) \text{ e } R(t)$$

dove

$$f(t) = -R'(t) = h(t) \cdot R(t) = 0.2t \cdot e^{-0.1t^2}$$

R(t) è la condizionante
h(t)dt è la condizionata

$$\text{con } dt = 0.01 \Rightarrow$$

$f(t)dt = h(t)dt$ se $R(t)=1$ ovvero è certa

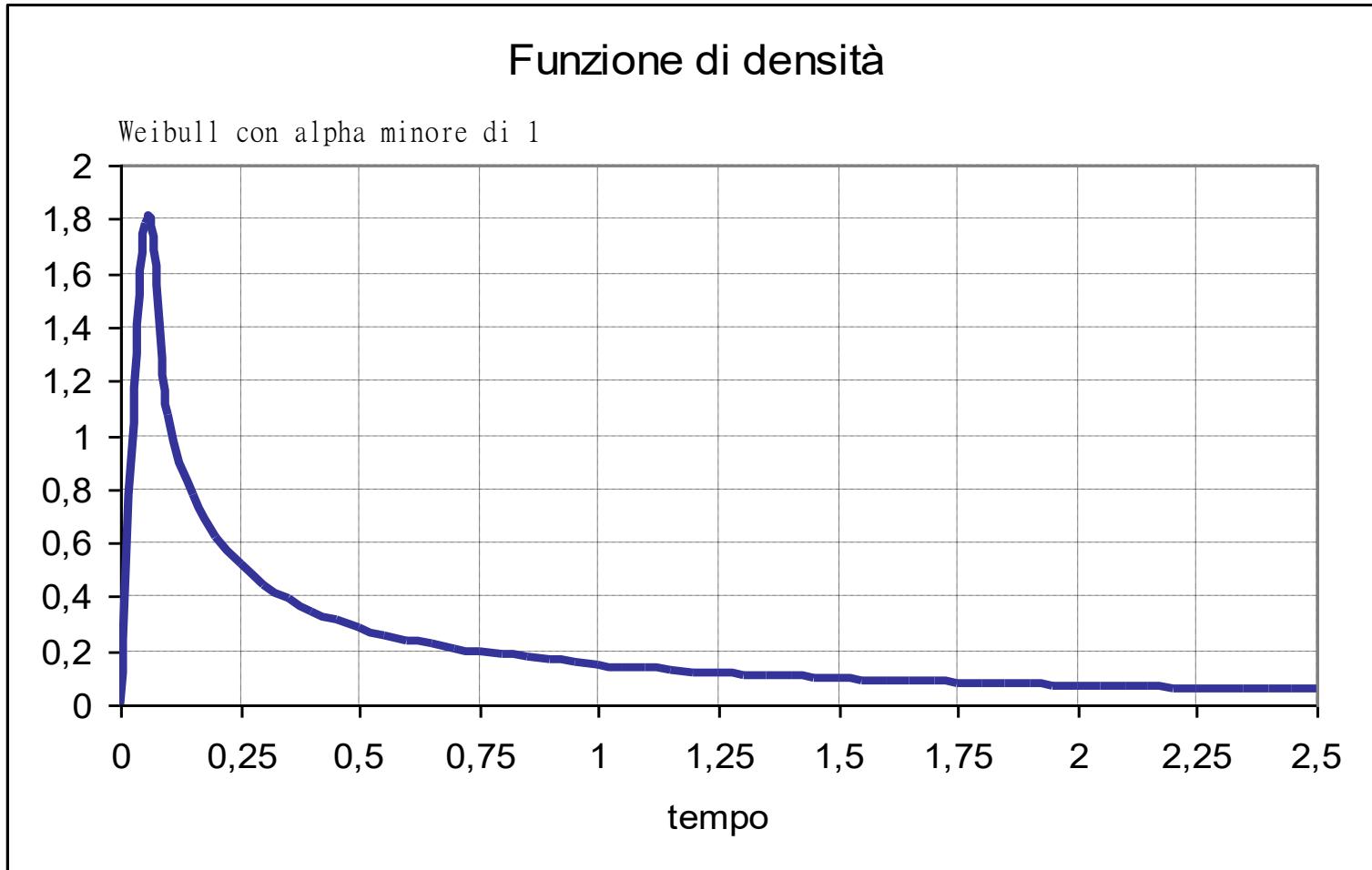
$$f(t)dt = f(0.25) \cdot 0.01 = 0.000497$$

Quantità
trascurabile

DIPENDE
DAL CONTESTO

Esercizio 1

(Considerazione sul quesito 3...)



Esercizio 2

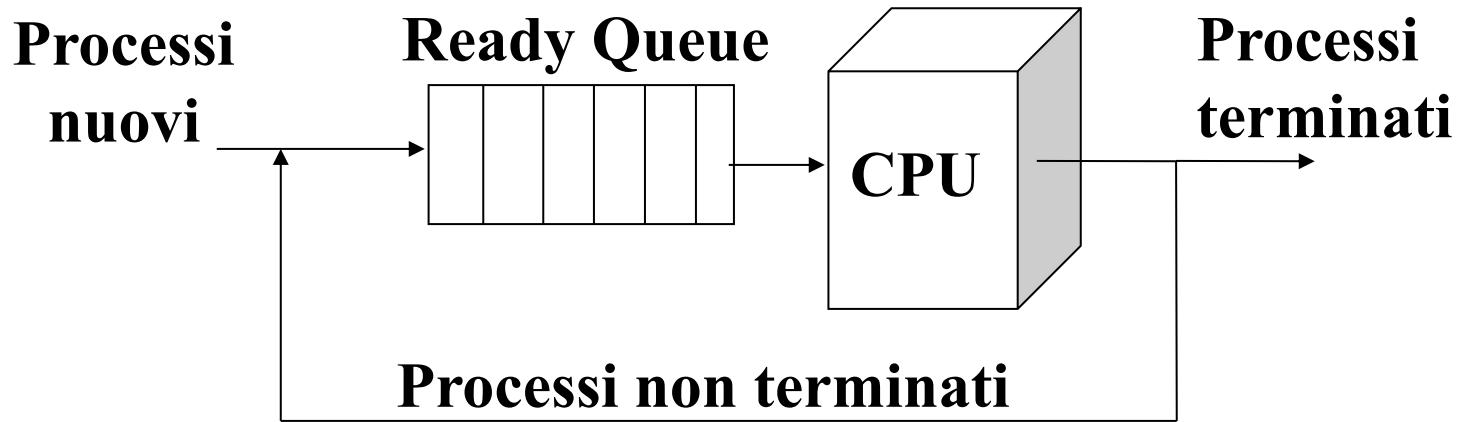
(Descrizione del problema)

Prova a riformularlo in un altro contesto

- Un sistemista deve stabilire la lunghezza del “*time slice*” per l’allocazione della CPU ai singoli processi in attesa nella “ready queue” di un sistema multiprogrammato.
- Al termine del *time slice* corrente il processo che non abbia ancora terminato l’esecuzione viene ricollocato nella “ready queue”, all’ultimo posto.
- Considerato che con un *monitor hardware* è stato determinato un tempo medio di esecuzione di 0,2sec per processo, il sistemista pensa di fissare a 0,02sec la durata del *time slice*.

Esercizio 2

(Descrizione del problema)



- Considerazioni?
- Ulteriori quesiti?

Esercizio 2 (Quesiti)

1. Proporre e giustificare un modello per il tempo di esecuzione di un processo. non si può usare Weibull perché si ha il tempo medio di esecuzione
2. Calcolare la probabilità che un processo abbia bisogno di più di un solo *time slice* per essere completamente eseguito.
3. Calcolare la probabilità che un processo abbia bisogno di un numero di *time slice* compresi tra 1 e 3, per essere completamente eseguito.
4. Posto che un processo sia stato rimesso in coda dopo aver esaurito il suo primo *time slice*, calcolare la probabilità che venga completato con il suo secondo *time slice*.

Esercizio 2 (Soluzione del quesito 1)

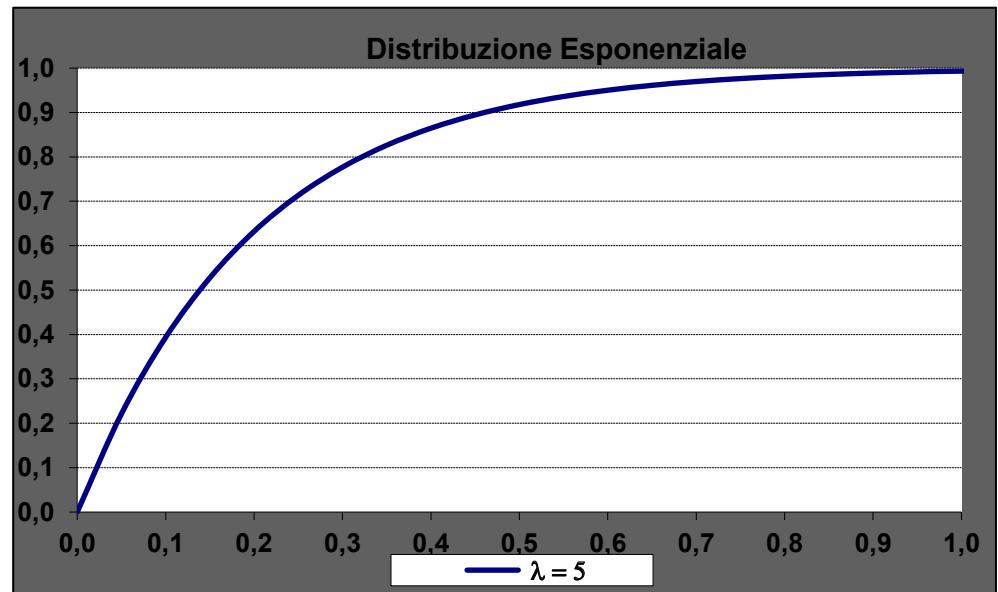
- Avendo solo la media (valore atteso) del tempo di esecuzione, il modello da adottare è quello esponenziale.

Sia X “la durata del tempo di esecuzione”

Allora: $E[X] = \frac{1}{\lambda} = \frac{2}{10} \Rightarrow \lambda = 5 \text{ sec}^{-1}$

e quindi

$$\begin{aligned} F_X(t) &= 1 - e^{-\lambda t} \\ &= 1 - e^{-5t} \end{aligned}$$



Esercizio 2 (Soluzione dei quesiti 2-4)

- La probabilità che sia sufficiente un unico *time slice* è:

$$P(X \leq 0.02) = \int_{0.0}^{0.2} f_X(u) du = F_X(0.02) = 0.095$$

- La probabilità che ne occorra più di uno solo è:

$$P(X > 0.02) = 1 - P(X \leq 0.02) = 0.905$$

- La probabilità che occorrono da 1 a 3 *time slice* è:

non uso il differenziale perché non è un tempo piccolo l'intervallo

$$P(0.02 \leq X < 0.06) = F_X(0.06) - F_X(0.02) = 0.164$$

- La probabilità che venga completato col suo secondo *time slice*, posto che non ce l'abbia fatta col suo primo.

$$P(X \leq 0.04 | X > 0.02) = P(X \leq 0.02) = 0.095$$

(Assenza di memoria)

Esercizio 2 (nuova soluzione del quesito 4)

SENZA RICORRERE ALLA PROPRIETA' DI ASSENZA DI MEMORIA

l'unica distribuzione che ha assenza di memoria è l'esponenziale

Sia: $Y \hat{=} X - \bar{x}$, dove: X è la “durata della vita del fenomeno”,
 Y la “vita residua” e x -segnato è l’età.

RISULTA:

$$P(Y \leq y \mid X \geq \bar{x}) = P(X - \bar{x} \leq y \mid X \geq \bar{x})$$

Vai a rivedere la dimostrazione alternativa

$$= \frac{P(\bar{x} \leq X \leq \bar{x} + y)}{P(X \geq \bar{x})} = \frac{F(\bar{x} + y) - F(\bar{x})}{1 - F(\bar{x})}$$

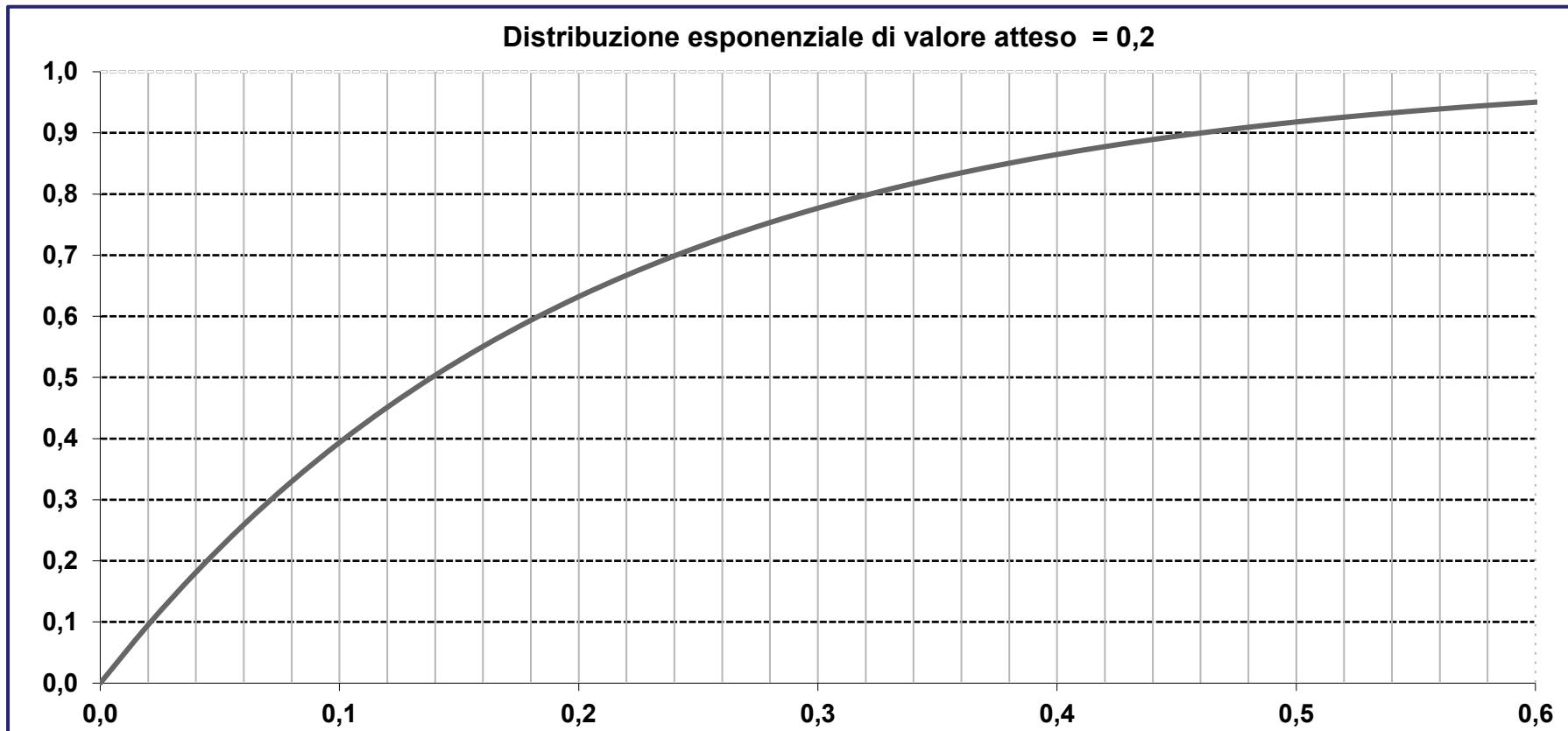
$$= \frac{F(0.02 + 0.02) - F(0.02)}{1 - F(0.02)} = 0.095$$

DOMANDA extra:

A cosa corrisponde la seguente?

$$\tilde{F}(x) \hat{=} \frac{F(x)}{F(\bar{x})}, \quad 0 \leq x \leq \bar{x}.$$

Esercizio 2 (verifica calcoli)



Nota: il 39% delle realizzazioni non supera la metà del valore atteso, ma il 5% lo supera di almeno 3 volte! (si usa il complementare)

Domanda extra: tracciare il grafico troncato al valore 0,6, cioè corrispondente ad un modello che escluda durate superiori di almeno 3 volte il valore atteso.

Esercizio 3

(la vista degradante con la vecchiaia)

vedi file per soluzione

Un professore vuole acquistare un paio di occhiali da vista, a causa della presbiopia che sta per arrivare a lui a causa dell'età.

Al momento il professore ci vede benissimo a leggere da vicino, diciamo al 100%.

Secondo il suo oculista, “la capacità visiva” (“ $C(t)$ ”) si ridurrà nel tempo in maniera aleatoria, con una velocità media costante. Precisamente, alla velocità media di 0,25 gradi all’anno per i primi 4 anni dopo l’acquisto.

E’ possibile calcolare la capacità visiva del professore dopo 2 anni, assumendo che sia pari a 10 gradi al momento dell’acquisto?

RICAP.: distribuzione della v.a. “max” (1_di_2)

- La v.a. “max” è:

$$X \doteq \max \{X_1, \dots, X_n\}$$

– Allora:

$$F_X(t) = Pr(X_1 \leq t, \dots, X_n \leq t) = \prod_{i=1}^n F_{X_i}(t)$$

– L'affidabilità aumenta assai poco con n :

$$F_X(t) = (1 - e^{-\lambda t})^n$$

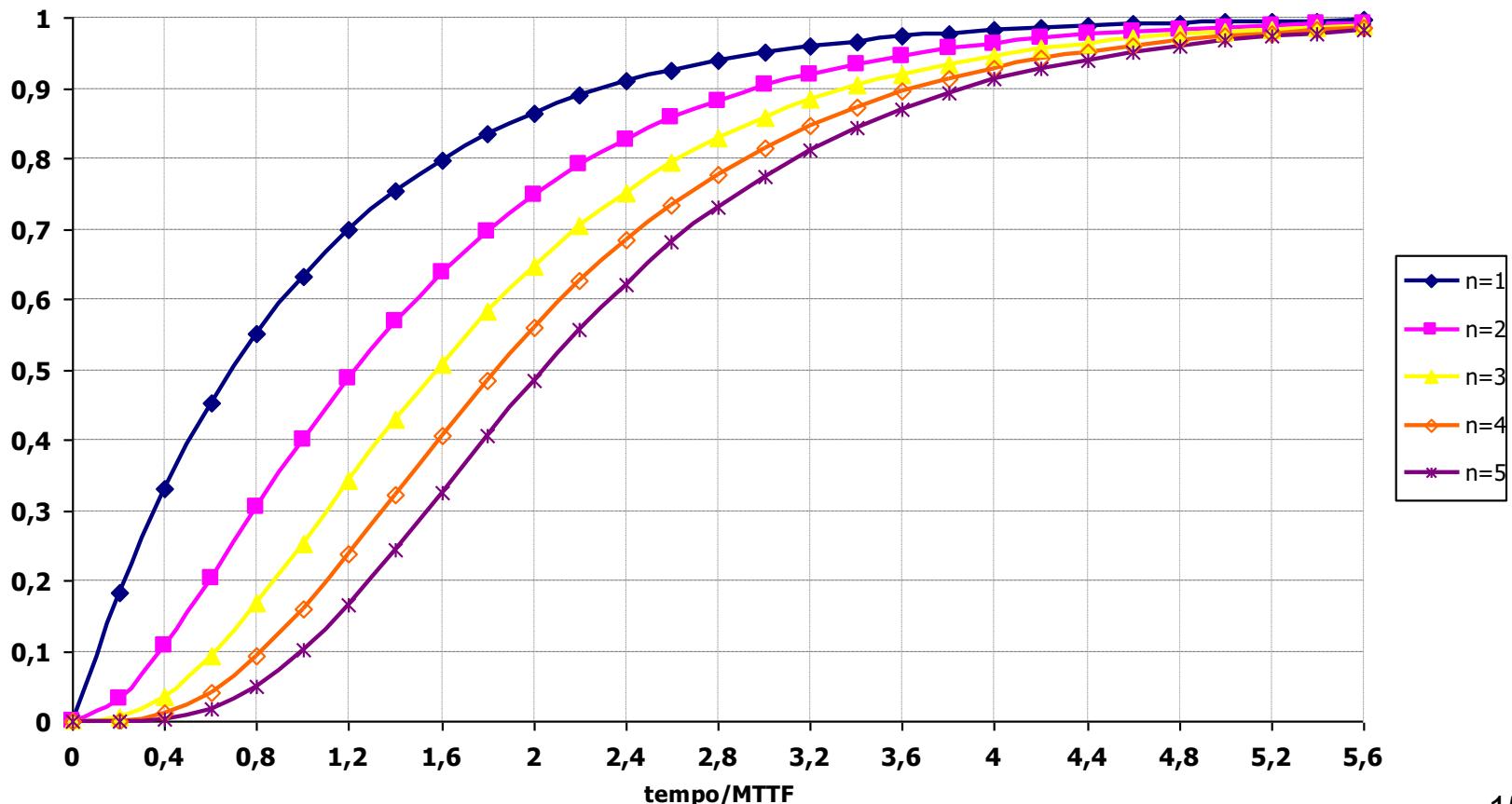
Caso
“esponenziale”

RICAP.: distribuzione della v.a. “max” (2_di_2)

Curve della $F_X(t)$ al variare di n

$$\lambda = 2$$

Funzione di Distribuzione v.a. "max"



RICAP.: distribuzione della v.a. “min” (1_di_2)

- La v.a. “min” è:

Allora:

$$X \doteq \min \{X_1, \dots, X_n\}$$

$$1 - F_X(t) = \Pr(X_1 > t, \dots, X_n > t) = \prod_{i=1}^n 1 - F_{X_i}(t)$$

L'affidabilità diminuisce assai velocemente con n :

$$F_X(t) = 1 - (e^{-\lambda t})^n$$

Caso
“esponenziale”

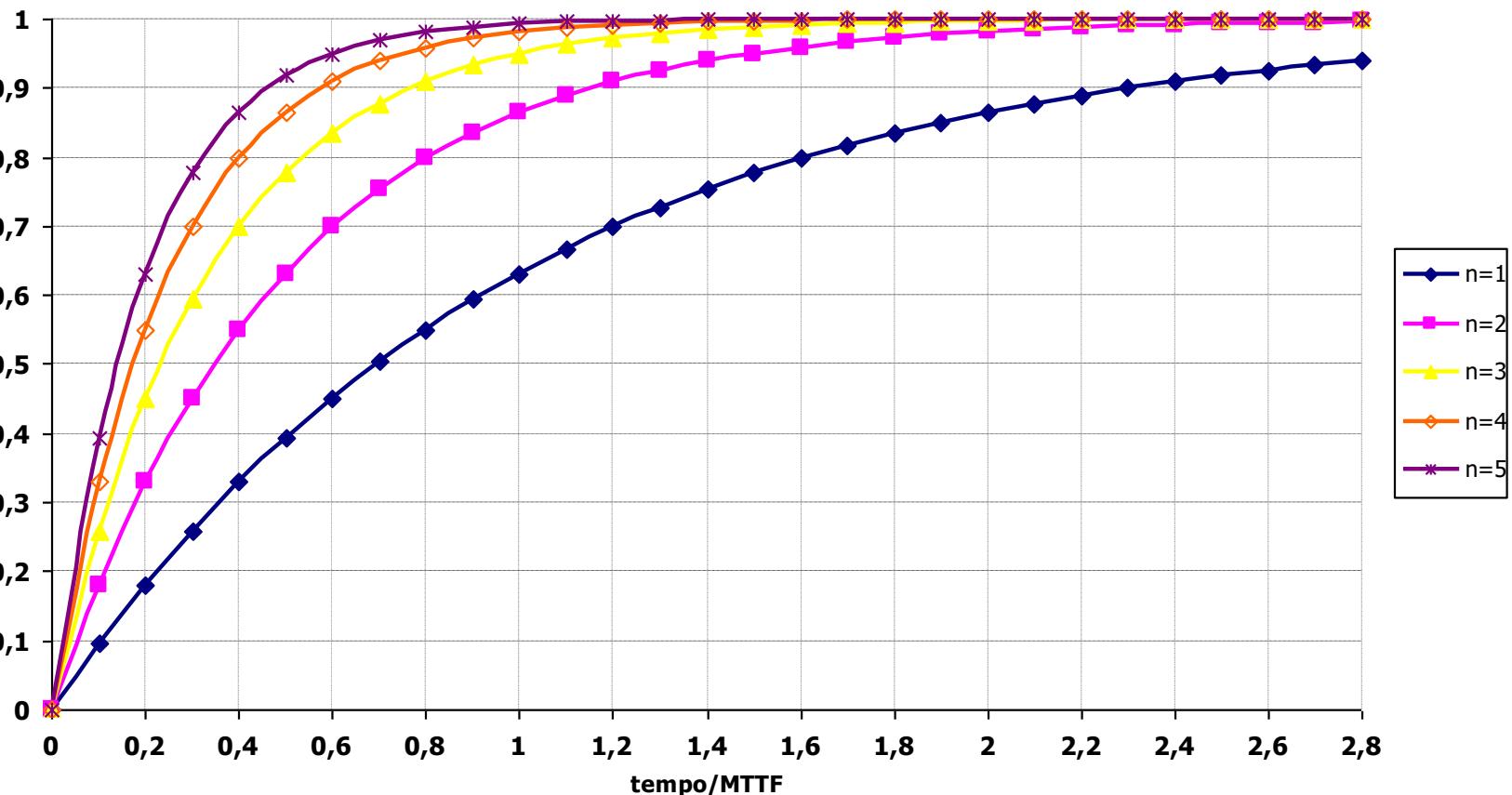
RICAP.: distribuzione della v.a. “min” (2_di_2)

Curve della $F_X(t)$ al variare di n

1

$$\lambda = 1$$

Funzione di Distribuzione v.a. "min"



RICAP.: generalizzazione per la v.a. “min di exp”

- Se

$$F_{X_i}(t) \hat{=} 1 - e^{-\lambda_i t}, \quad i = 1, \dots, n$$

e indipendenti

$$X \hat{=} \min \{X_1, \dots, X_n\}$$

allora la X segue ancora una legge esponenziale:

$$F_X(t) = 1 - \prod_{i=1}^n e^{-\lambda_i t} = 1 - \exp[-t \sum_{i=1}^n \lambda_i]$$

Esercizio 4

(Descrizione del problema)

Classica situazione per evitare la consistenza

- Un certo numero di processi-lettori può accedere alle tabelle di un database.
- Due processi-scrittori possono aggiornare il DB lavorando in **parallelo** (su tabelle diverse) e in modo **indipendente**.
- L'accesso dei processi-lettori al DB è bloccato per tutta la durata dell'aggiornamento.

Esercizio 4 (Quesiti)

- Adottando due modelli esponenziali, di media pari a 4 sec e 6 sec, per le rispettive durate del lavoro dei due processi-scrittori:
 1. Ricavare la distribuzione della v.a.
 X = “tempo di blocco degli accessi in lettura”.
(quanto tempo deve aspettare un processo per leggere?) il massimo fra i due tempi di esecuzione dei processi scrittori
 2. Calcolarne il valore atteso.

Esercizio 4 (Soluzione al quesito 1)

- Risulta:

$$X \doteq \max \{X_1, X_2\}$$

- dove X_i rappresenta il contributo alla durata del blocco dovuto al processo i (con $i=1,2$).
- Allora, usando due leggi exp di parametro μ_1 e μ_2 :

Utilizzo affidabilità nel caso esponenziale (vedi corollario)

$$\begin{aligned} F_X(t) &= (1 - e^{-\mu_1 t})(1 - e^{-\mu_2 t}) \\ &= 1 - (e^{-\mu_1 t} + e^{-\mu_2 t} - e^{-(\mu_1 + \mu_2)t}) \end{aligned}$$

Esercizio 4

(Soluzione al quesito 2)

CALCOLO

- Il valore atteso della durata del blocco:

$$\begin{aligned} E[X] &= \int_{t=0}^{\infty} (1 - F_X(t)) dt \quad \text{che tra l'altro è l'MTTF. Appunto integrale della sopravvivenza} \\ &= \int_{t=0}^{\infty} \left(e^{-\mu_1 t} + e^{-\mu_2 t} - e^{-(\mu_1 t + \mu_2 t)} \right) dt \\ &= \frac{1}{\mu_1} + \frac{1}{\mu_2} - \frac{1}{\mu_1 + \mu_2} \cong 7.5 \text{ sec} \end{aligned}$$

cos'è seconde te?

Esercizio 5

(Descrizione del problema)

- Due processi indipendenti elaborano dati letti dalle tabelle di un database e dopo aggiornano le tabelle.
- Quello che completa per **primo l'elaborazione** procede con l'operazione di **aggiornamento**.
- La **durata media** delle elaborazioni effettuate dai due processi è pari a 2 sec e 4 sec, rispettivamente.

Esercizio 5

Quesiti

- Adottando due modelli esponenziali, per le durate delle elaborazioni dei due processi:
 1. Ricavare la distribuzione della v.a.
 X = “tempo al prossimo aggiornamento”.
il minimo che serve per il prossimo aggiornamento
 2. Calcolarne il valore atteso.

Esercizio 5 (Soluzione)

- Risulta:

$$X \doteq \min \{X_1, X_2\}$$

- dove X_i rappresenta la durata dell'elaborazione effettuata dal processo i (con $i=1,2$).

utilizzando la definizione di \min vista pure nella ricapitolazione

- Allora, usando due leggi \exp di parametro μ_1 e μ_2 :

$$1 - F_X(t) = e^{-\mu_1 t} \cdot e^{-\mu_2 t}$$

$$F_X(t) = 1 - e^{-(\mu_1 + \mu_2)t}; \quad E[X] = \frac{1}{\mu_1 + \mu_2} = 1.33 \text{ sec}$$

valore atteso

Richiamo formule sistemi “m-out-of-n”

- Durate indipendenti e identiche, con stessa legge esponenziale e durata media = $1/\lambda$:

$$R_{1/n}^{\text{parallelo}}(t) = \sum_{j=1}^n \binom{n}{j} \left[e^{-\lambda t} \right]^j \left[1 - e^{-\lambda t} \right]^{n-j} = 1 - \left[1 - e^{-\lambda t} \right]^n$$

$$R_{m/n}(t) = \sum_{j=m}^n \binom{n}{j} \left[e^{-\lambda t} \right]^j \left[1 - e^{-\lambda t} \right]^{n-j}, \quad 1 < m < n$$

$$R_{n/n}^{\text{seriale}}(t) = \sum_{j=n}^n \binom{n}{j} \left[e^{-\lambda t} \right]^j \left[1 - e^{-\lambda t} \right]^{n-j} = \left[e^{-\lambda t} \right]^n$$

Esercizio 6

(il problema dell'attesa in pizzeria) (1_di_2)

È martedì sera; una piccola pizzeria apre alle ore 20 e 30 e tutti gli 8 tavoli disponibili sono prenotati esattamente per quell'ora. Una coppia di amici arriva proprio alle 20 e 30, però senza prenotazione.

Gli 8 tavoli prenotati sono tutti occupati da studenti universitari, che hanno comportamenti identici dal punto di vista probabilistico. Pertanto, si può assumere che il tempo medio di occupazione del generico tavolo sia di 60 min. Identici dal punto di vista probabilistico è diverso da dire uguale

Inoltre, non ci sono due o più tavoli occupati da uno stesso gruppo. Gli studenti seduti a tavoli diversi neppure di vista si conoscono!

Vuol dire che sono identicamente distribuiti
ma indipendenti!

Ipotizzando che un modello probabilistico con assenza di memoria per la variabile aleatoria “tempo di occupazione di uno qualunque dei tavoli” sia adeguato, si consideri pure trascurabile la probabilità che due o più tavoli si liberino nello stesso istante (per puro caso).

che modello si può usare? quale no?

La porzione di testo in blu è una ipotesi semplificativa
che varrà sempre ed è molto semplificativa

Esercizio 6

(il problema dell'attesa in pizzeria) (2_di_2)

Nelle ipotesi fatte,

- 1) calcolare la probabilità che la coppia di amici possa avere un tavolo entro 20 minuti. sarebbe il min

MODIFICA del problema

La stessa pizzeria non ha un orario di chiusura prestabilito e ha accettato solo il primo turno di prenotazioni, mandando via la coppia di amici senza prenotazione e chiunque altro arrivasse dopo. In tale nuova ipotesi,

- 2) calcolare la probabilità che il gestore possa chiudere entro 120 minuti dalle 20 e 30 sarebbe il calcolo del max

Lo stesso gestore ha due camerieri in servizio, ma ha deciso che ne manderà via uno alle 22 e 30 se per quell'ora saranno ancora occupati uno o due tavoli (sugli otto occupati dalle 20 e 30).

- 3) Calcolare la probabilità che ciò accada. m out of n caso M=1 e caso M=2

Data l'ipotesi di utilizzare un modello probabilistico con assenza di memoria, risulta naturale utilizzare il modello esponenziale: $F(t) = 1 - e^{-\lambda t}$

Il numero di tavoli totali è 8 e come suggerisce il testo, il tempo medio di occupazione di un tavolo è di 60 minuti. Dunque il valore atteso è 60 min e da esso si può ricavare lambda: $E[X] = \frac{1}{\lambda} = 60 \text{ min} \implies \lambda = \frac{1}{E[X]} = \frac{1}{60 \text{ min}} = 0.017 \text{ min}^{-1} = 0.00027778 \text{ s}^{-1}$

Risoluzione quesito 1): un tavolo qualunque si deve liberare entro 20 min

Consideriamo gli 8 tavoli come un sistema seriale. L'affidabilità è dunque pari a $R_{ser}(t)$ e prevede che nessun tavolo si liberi prima di un tempo t. Con $1-R_{ser}(t)$ calcolo la probabilità che almeno uno dei tavoli si liberi. Cioè quando un tavolo smette di essere occupato il sistema non è più "affidabile", se fossero pistone di una macchina sarebbe quando uno qualunque su 8 si rompe e va sostituito no?

Usando il risultato che esprime l'affidabilità per un sistema seriale si ottiene che:

$$R_{ser}(t) = 1 - F_{Y_1}(t) = 1 - (1 - \exp[-\sum_{i=1}^n \lambda_i]t) = \exp[-n\lambda t] = e^{-8 \cdot 0.00027778 \cdot 20 \cdot 60} = 0.069$$

In particolare se X="tempo che un tavolo si liberi"

$$\Pr(X \leq 20) = 1 - R_{ser}(t) = 1 - 0.069 = 0.931$$

Risoluzione quesito 2): ovvero tutti i tavoli si devono liberare entro 120 min

Stavolta considero gli 8 tavoli come un sistema parallelo. L'affidabilità in questo caso è pari a $R_{par}(t)$ e prevede che almeno 1 dei tavoli non si liberi prima di un tempo t. Con $1-R_{par}(t)$ calcolo la probabilità che tutti i tavoli si liberino entro un tempo t.

Usando il risultato che esprime l'affidabilità per un sistema parallelo si ottiene che:

$$R_{par}(t) = 1 - [1 - R(t)]^n = 1 - (F_X(t))^n = 1(1 - e^{-\lambda t})^n = 1 - (1 - e^{-0.00027778 \cdot 120 \cdot 60})^8 = 0.68$$

In particolare se X="tempo che tutti i tavoli siano liberi"

$$\Pr(X \leq 120) = 1 - R_{par}(t) = 1 - 0.68 = 0.32$$

Risoluzione quesito 3): max 2 tavoli occupati entro 2 ore perché il servizio parte alle 20:30

Considero un sistema 3 out of 8. L'affidabilità di questo sistema prevede che almeno 3 tavoli su 8 siano occupati al tempo t. Con $1-R_{3/8}(t)$ calcolo la probabilità che MENO DI 3 tavoli siano occupati al tempo t.

Esercizi di Riepilogo

Esercizio 1

Il responsabile di produzione di un'azienda manifatturiera deve stabilire la lunghezza del “time slice” per l’allocazione temporanea di una macchina (condivisa) ai singoli semilavorati presenti in un’area di stoccaggio vicina alla macchina stessa. In pratica, la macchina preleva un semilavorato e lo lavora per un certo tempo (time slice). Se la lavorazione del semilavorato rimane incompleta alla fine del time slice, allora la macchina rideposita lo stesso semilavorato nell’area di stoccaggio e lo riprenderà più tardi; altrimenti, il semilavorato lascia il reparto come prodotto finito. La ratio di questa politica consiste nel voler dare priorità ai semilavorati che avrebbero bisogno di meno tempo per essere trasformati in prodotti finiti, pur senza conoscere i tempi richiesti da ognuno in anticipo, cioè al momento dell’allocazione della macchina al semilavorato.

Considerato che la durata media della lavorazione necessaria a realizzare il prodotto finito è pari a 2 ore e che il time slice è stato fissato pari al 25% di detta durata media:

1. proporre e giustificare un modello probabilistico per il tempo di lavorazione del semilavorato;
2. calcolare la probabilità che un semilavorato abbia bisogno di più di un time slice per essere completato;
3. calcolare la probabilità che un semilavorato abbia bisogno di un numero di time slice compreso tra 1 e 3 per essere completato;
4. posto che un semilavorato sia stato ridepositato nell’area di stoccaggio temporaneo dopo il primo time slice di lavorazione, calcolare la probabilità che lo stesso venga completato col prossimo time slice in cui gli verrà riassegnata la macchina;
5. posto che un semilavorato sia stato ridepositato nell’area di stoccaggio temporaneo dopo il primo time slice di lavorazione, proporre un modello probabilistico per calcolare la probabilità che lo stesso venga ripreso dopo 1, 2, 3, ... n time slice.
6. assumendo che occorrono j time slice ($j = 1, 2, \dots, k$) per completare un semilavorato, calcolare la probabilità che esso venga completato entro gli n time slice di cui è composto un turno di lavoro;

7. infine, riproporre un modello probabilistico alternativo a quello usato fino a questo punto che sia capace di escludere la possibilità che si verifichino realizzazioni troppo piccole (vicine allo zero, infinitesimali) pur riuscendo a contemplare una significativa probabilità di realizzazioni relativamente piccole e pure una probabilità non trascurabile di realizzazioni estremamente grandi.

Soluzione

- Il modello probabilistico da proporre deve avere la caratteristica di riprodurre un alto numero di tempi piccoli (per i semilavorati che avrebbero bisogno di meno tempo per essere trasformati in prodotti finiti). Al tempo stesso, si dispone del solo valore medio. Quindi, il modello probabilistico per il tempo di lavorazione è la funzione esponenziale.

Considerando la durata media di lavorazione di 2 ore, il time slice fissato pari al 25% ovvero 0,5 ore, se si definisce la variabile aleatoria $X \hat{=} \text{tempo di lavorazione}$ allora il modello è:

$$F_X(t) = 1 - e^{-\lambda t} \text{ con } \lambda = 0,5 h^{-1}.$$

- La probabilità che un semilavorato abbia bisogno di più di un time slice per essere completato è:

$$P(X > 0,5) = 1 - P(X \leq 0,5) = 1 - F_X(0,5) = 1 - 1 + e^{-0,5 \cdot 0,5} \cong 0,78.$$

- La probabilità che un semilavorato abbia bisogno di un numero di time slice compreso tra 1 e 3 per essere completato è:

$$P(0,5 \leq X \leq 1,5) = P(X \leq 1,5) - P(X \leq 0,5) = F_X(1,5) - F_X(0,5) = 1 - e^{-0,5 \cdot 1,5} - 1 + e^{-0,5 \cdot 0,5} = -0,47 + 0,78 = 0,31.$$

- Posto che un semilavorato sia stato ridepositato nell'area di stoccaggio temporaneo dopo il primo time slice di lavorazione, per calcolare la probabilità che lo stesso venga completato col prossimo time slice in cui gli verrà riassegnata la macchina si ricorre alla formula della probabilità condizionata:

$$P(X \leq 1 | X > 0,5) = \frac{P(0,5 < X \leq 1)}{P(X > 0,5)} = \frac{P(X \leq 1) - P(X \leq 0,5)}{1 - P(X \leq 0,5)} = \frac{F_X(1) - F_X(0,5)}{1 - F_X(0,5)} =$$

$$\frac{1 - e^{-0,5 \cdot 1} - 1 + e^{-0,5 \cdot 0,5}}{1 - 1 - e^{-0,5 \cdot 0,5}} = \frac{-0,61 + 0,78}{0,78} \cong 0,22.$$

È possibile pervenire allo stesso risultato sfruttando la proprietà di assenza della memoria della funzione di distribuzione esponenziale:

$$P(X \leq 1 | X > 0,5) = P(X \leq 0,5) = F_X(0,5) = 1 - e^{-0,5 \cdot 0,5} \cong 1 - 0,78 = 0,22.$$

5. Se si codifica come successo la ripresa di un semilavorato per la lavorazione, se questo viene ripreso dopo 1,2,3,...,n time slice ciò significa che si ottiene un successo al tentativo 1,2,3,...,n. Si può quindi proporre la legge geometrica come modello probabilistico per calcolare la probabilità che lo stesso venga ripreso dopo 1,2,3,...,n time slice:

$$P(S = K | \infty) = q^{k-1} \cdot p.$$

6. Se si codifica come successo l'assegnazione di un time slice per la lavorazione allora ciò significa che occorrono j successi ($j = 1,2,\dots,k$), non necessariamente consecutivi, entro gli n time slice del turno di lavoro. Si può quindi proporre la legge binomiale come modello probabilistico per calcolare la probabilità che esso venga completato entro gli n time slice di cui è composto un turno di lavoro:

$$P(K = j | n) = \binom{n}{j} p^j \cdot q^{n-j} \text{ con } (j = 1,2,\dots,k).$$

7. Il modello di Weibull è capace di escludere la possibilità che si verifichino realizzazioni troppo piccole (vicine allo zero, infinitesimali) e, al tempo stesso, contemplare un significativo numero di realizzazioni non nullo in corrispondenza di valori molto grandi. L'esponenziale, al contrario, è caratterizzato da un grande numero di realizzazioni prossimi allo zero:

$$f_X(t) = \lambda \alpha t^{\alpha-1} e^{-\lambda t^\alpha},$$

$F_X(t) = 1 - e^{-\lambda^{\alpha}}$ con $\alpha > 0$, $\lambda > 0$ (se $0 < \alpha < 1$ si ha un decreasing failure rate, $\alpha = 1$ un constant failure rate e $\alpha > 1$ un increasing failure rate).

Esercizio 2

Un sistema elettrico dispone di un componente elettrico non riparabile, ma funzionante all'origine. Detto componente non è soggetto ad usura nel tempo.

Per esso:

1. proporre la funzione di affidabilità coerente con il modello di usura descritto;
2. calcolare il tempo medio al guasto;
3. calcolare la probabilità che si guasti entro la vita media;
4. calcolare la probabilità che si guasti tra la vita media e il doppio di essa;
5. calcolare il rischio di guasto in corrispondenza della vita media, così come risulterebbe all'istante di messa in opera del componente;
6. calcolare la probabilità del guasto in un intervallino di ampiezza pari alla vita media più l'1% di essa, così come risulterebbe all'istante di messa in opera del componente;
7. calcolare la probabilità di guasto in corrispondenza di un'età pari alla vita media più l'1% di essa;
8. calcolare la probabilità che il componente subisca il guasto tra due istanti di tempo corrispondenti, rispettivamente, alla vita media e al doppio di essa, dando per scontato che ad avere un'età pari alla vita media ci arrivi.

Soluzione

1. Il componente elettrico del sistema in esame non è soggetto ad usura nel tempo ovvero il suo tasso di guasto è costante $h(t) = c = \lambda$. Per definire una funzione di affidabilità coerente con il modello di usura descritto, si può ricorrere alla relazione che lega il tasso di guasto del componente alla funzione di affidabilità stessa, cioè:

$$R(u) = \exp \left[- \int_{t=0}^u h(t) dt \right] = \exp \left[- \int_{t=0}^u c dt \right] = \exp \left[- \lambda t \Big|_0^u \right] = e^{-\lambda u}.$$

2. Indicando con X la variabile aleatoria “tempo al guasto”, per calcolare il tempo medio al guasto del componente detto anche $MTTF$ – *mean time to failure* si può utilizzare la formula alternativa del calcolo del valore atteso:

$$E[X] = \int_{x=0}^{\infty} [1 - F_X(x)] dx = \int_{x=0}^{\infty} R(x) dx = \int_{x=0}^{\infty} e^{-\lambda x} dx = -\frac{1}{\lambda} e^{-\lambda x} \Big|_0^{\infty} = \frac{1}{\lambda},$$

quindi $MTTF = \frac{1}{\lambda}$.

3. La probabilità che il componente si guasti entro la vita media è data da:

$$P(X \leq MMTF) = F_X(MTTF) = 1 - e^{-\lambda \frac{1}{\lambda}} \cong 1 - 0,37 = 0,63.$$

4. La probabilità che il componente si guasti tra la vita media ed il doppio di essa è data da:

$$\begin{aligned} P(MTTF \leq X \leq 2 \cdot MMTF) &= F_X(2 \cdot MTTF) - F_X(MTTF) = 1 - e^{-\lambda \cdot 2 \cdot \frac{1}{\lambda}} - 1 + e^{-\lambda \cdot \frac{1}{\lambda}} \\ &\cong -0,14 + 0,37 = 0,23 \end{aligned}$$

5. Il rischio di guasto in corrispondenza della vita media, così come risulterebbe all’istante di messa in opera del componente, è dato da:

$$h(t) = \frac{f_x(MTTF)}{1 - F_x(MTTF)} = \frac{\lambda e^{-\lambda \frac{1}{\lambda}}}{1 - 1 + e^{-\lambda \frac{1}{\lambda}}} = \frac{\lambda \cdot 0,37}{0,37} = \lambda.$$

6. La probabilità del guasto in un intervallino di ampiezza pari alla vita media più l’1% di essa, così come risulterebbe all’istante di messa in opera del componente, non è altro che la probabilità che il componente si guasti tra $t, t + dt$ dal momento che l’1% può essere considerato un dt di $MTTF$, ovvero:

$$f_x(t) \cdot dt = \lambda e^{-\lambda \frac{1}{\lambda}} \cdot 0,01 \cdot MTTF = \lambda e^{-\lambda \frac{1}{\lambda}} \cdot 0,01 \cdot \frac{1}{\lambda} \cong 0,01 \cdot 0,37 = 0,0037.$$

Osservazione importante: in base a questa impostazione il componente non ha alcuna età.

7. La probabilità di guasto in corrispondenza di un'età pari alla vita media più l'1% di essa è data da:

$$h(t) \cdot dt = \lambda \cdot 0,01 \cdot \lambda = 0,01 \cdot \lambda^2.$$

Osservazione importante: questa impostazione è basata sul fatto di sapere che il componente ha un'età.

8. La probabilità che il componente subisca il guasto tra due istanti di tempo corrispondenti, rispettivamente, alla vita media e al doppio di essa, dando per scontato che ad avere un'età pari alla vita media ci arrivi:

$$\begin{aligned} P(X \leq 2 \cdot MMTF \mid X > MTTF) &= \frac{P(MTTF < X \leq 2 \cdot MMTF)}{P(X > MMTF)} \\ &= \frac{P(X \leq 2 \cdot MMTF) - P(X \leq MMTF)}{1 - P(X \leq MMTF)} \\ &= \frac{F_X(2 \cdot MMTF) - F_X(MMTF)}{1 - F_X(MMTF)} = \frac{1 - e^{-\lambda \cdot 2 \cdot \frac{1}{\lambda}} - 1 + e^{-\lambda \cdot \frac{1}{\lambda}}}{1 - 1 + e^{-\lambda \cdot \frac{1}{\lambda}}} \\ &\approx \frac{0,86 - 0,63}{0,37} = 0,62. \end{aligned}$$

Esercizio 3

Lungo la banchina del porto di Gioia Tauro ha appena ormeggiato una nave portacontainer. Per le operazioni di scarico e carico è previsto l'impiego, in parallelo, di 2 gru identiche, mobili su rotaia. Le gru stanno per arrivare sul punto di ormeggio occupato dalla nave e le operazioni di scarico potranno cominciare non appena sarà arrivata la prima delle 2 gru in movimento.

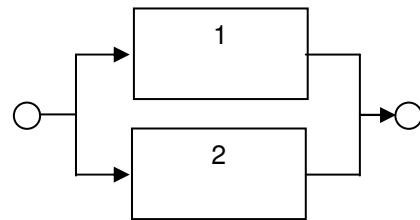
1. Ipotizzando una durata media identica e pari a 10 minuti, per il movimento in atto di entrambe le gru, proporre un modello per la variabile aleatoria “tempo all’arrivo della prima gru” e con esso calcolare la probabilità che occorrono più di 5 minuti per avviare le operazioni di scarico.
2. Una volta avviate, le operazioni di scarico di ciascuna gru durano in media 4 ore e la durata può essere riferita ad uno stesso modello esponenziale. Calcolare la

probabilità che le operazioni complessive sulla nave vengano completate entro 4 ore dall'istante d'arrivo della seconda gru, che è possibile fissare come istante “0”.

3. Una porzione di stiva di una certa nave contiene “n=10” alloggiamenti per container, ognuno dei quali ha probabilità pari a “p=0.25” di essere occupato. Proporre due “modelli di occupazione” della porzione di stiva e con essi valutare la probabilità che almeno k=2 container siano presenti.

Soluzione

Le 2 gru sono identiche e lavorano in parallelo



1. Definite le seguenti variabili aleatorie

$X \hat{=} \text{tempo all'arrivo della prima gru}$

$X_1 \hat{=} \text{tempo all'arrivo della gru 1}, F_{X_1}(t) = 1 - e^{-\lambda_1 t}$

$X_2 \hat{=} \text{tempo all'arrivo della gru 2}, F_{X_2}(t) = 1 - e^{-\lambda_2 t} \text{ con } \lambda_1 = \lambda_2 = \frac{1}{10} \text{ minuti}^{-1}$,

X è di fatto definita dalla gru che arriva per prima ovvero $X = \min\{X_1, X_2\}$. Il minimo di due funzioni esponenziali è ancora una funzione esponenziale con esponente pari alla somma dei singoli esponenti:

$$F_X(t) = 1 - e^{-(\lambda_1 + \lambda_2)t} = 1 - e^{-2\lambda t}.$$

La probabilità che occorrono più di 5 minuti per avviare le operazioni di scarico è, quindi, data da:

$$P(X > 5) = 1 - P(X \leq 5) = 1 - F_X(5) = 1 - 1 + e^{-2 \cdot \frac{1}{10} \cdot 5} = e^{-1} \cong 0,37$$

2. Definite le seguenti variabili aleatorie

$Y \hat{=} \text{durata complessiva delle operazioni di carico/scarico sulla nave}$

$Y_1 \hat{=} \text{durata delle operazioni effettuate dalla gru 1}, F_{Y_1}(t) = 1 - e^{-1 \cdot t}$

$Y_2 \hat{=} \text{durata delle operazioni effettuate dalla gru 2}, F_{Y_2}(t) = 1 - e^{-2 \cdot t} \text{ con}$

$$_1 = _2 = \frac{1}{4} \text{ ore}^{-1},$$

Y è di fatto definita dalla gru che termina per ultima ovvero $Y = \min\{Y_1, Y_2\}$:

$$\begin{aligned} F_Y(t) &= F_{Y_1}(t) \cdot F_{Y_2}(t) \\ &= (1 - e^{-1 \cdot t}) \cdot (1 - e^{-2 \cdot t}) = (1 - e^{-1 \cdot t}) \cdot (1 - e^{-2 \cdot t}) \\ &= 1 - e^{-1 \cdot t} - e^{-2 \cdot t} + e^{-2 \cdot t} = 1 - 2e^{-1 \cdot t} + e^{-2 \cdot t} \end{aligned}$$

La probabilità che le operazioni complessive sulla nave vengano completate entro 4 ore è, quindi, data da:

$$\begin{aligned} P(X \leq 4) &= 1 - F_Y(4) = 1 - 2e^{-\frac{1}{4}} + e^{-\frac{2}{4}} = 1 - 2e^{-1} + e^{-2} \\ &\approx 1 - 0,736 + 0,135 \\ &= 0,399 \end{aligned}$$

3. Indicando come “successo” lo stato di “alloggiamento stiva occupato”, allora la probabilità che almeno $k = 2$ container siamo presenti nella stiva che contiene un numero di alloggiamenti $n = 10$ può esprimersi mediante la formula della probabilità binomiale con $n = 10$ e $k \geq 2$ oppure “ m out of n ” se si vede come “funzionante” lo stato di “alloggiamento stiva occupato”.

Essendo $P(\text{successo}) = p = 0,25$, allora $q = 0,75$. Quindi,

$$\begin{aligned} P(\text{almeno 2 container presenti}) &= \sum_{k=2}^{10} \binom{n}{k} p^k \cdot q^{n-k} \\ &= \binom{10}{2} 0,25^2 \cdot 0,75^8 + \binom{10}{3} 0,25^3 \cdot 0,75^7 + \binom{10}{4} 0,25^4 \cdot 0,75^6 \\ &\quad + \binom{10}{5} 0,25^5 \cdot 0,75^5 + \binom{10}{6} 0,25^6 \cdot 0,75^4 + \binom{10}{7} 0,25^7 \cdot 0,75^3 \\ &\quad + \binom{10}{8} 0,25^8 \cdot 0,75^2 + \binom{10}{9} 0,25^9 \cdot 0,75 + \binom{10}{10} 0,25^{10} \end{aligned}$$

$$\begin{aligned}
&= 45 \cdot 0,25^2 \cdot 0,75^8 + 120 \cdot 0,25^3 \cdot 0,75^7 + 210 \cdot 0,25^4 \cdot 0,75^6 \\
&\quad + 252 \cdot 0,25^5 \cdot 0,75^5 + 210 \cdot 0,25^6 \cdot 0,75^4 + 120 \cdot 0,25^7 \cdot 0,75^3 \\
&\quad + 45 \cdot 0,25^8 \cdot 0,75^2 + 10 \cdot 0,25^9 \cdot 0,75 + 0,25^{10} \\
&= 0,28 + 0,25 + 0,15 + 0,06 + 0,02 + 0,003 \\
&\quad + 0,0004 + 0,00003 + 0,000001 \\
&\cong 0,76
\end{aligned}$$

Esercizio 4

L'elaborazione di alcuni dati acquisiti da un aeromobile durante il volo viene fatta da quattro processori indipendenti che lavorano in parallelo.

Ad ognuno viene dato lo stesso input e si suppone che i risultati siano corretti se almeno *tre dei quattro* processori restituiscono lo stesso risultato.

Il comparatore dei risultati può essere considerato "perfetto" e la legge di guasto, identica per i quattro componenti, è quella esponenziale, con un MTTF pari a *10.000 ore*:

1. Calcolare la probabilità che nessuno dei quattro componenti si guasti prima del 5% del proprio MTTF.
2. Qual è la probabilità che il sistema superi le *500 ore di vita*? [n.b. il sistema funziona se funzionano almeno 3 dei 4 componenti]
3. Qual è, invece, la probabilità che dopo le *500 ore di vita* risultino funzionanti *non più di due dei quattro* componenti?
4. Calcolare il tasso di guasto del sistema alle *500 ore di vita*.
5. Ricalcolare il tasso di guasto del sistema alle *500 ore di vita* supponendo, però, che alle 500 ore il sistema si sia già ridotto a 3 componenti, per effetto di uno ed un solo guasto.

Soluzione

1. Sia X_i il tempo al guasto del componente i -esimo ($i = 1, \dots, 4$).

Dal testo, si sa che $\lambda_i = \frac{1}{10000} = 10^{-4}$ ore⁻¹ $\forall i = 1, \dots, 4$, ossia, $\lambda_i = \lambda$, $i = 1, \dots, 4$.

Pertanto le funzioni di distribuzione saranno $F_{X_i}(t) = 1 - e^{-\lambda t}$.

Definita la variabile aleatoria

$$X = \min\{X_1, X_2, X_3, X_4\},$$

si può calcolare la probabilità che nessuno dei quattro componenti si guasti prima del 5% del proprio MTTF come:

$$P(X > t) = 1 - F_X(t) = \exp\left[-t \sum_{i=1}^4 \lambda_i\right].$$

Per sostituzione dei valori noti si ottiene:

$$P(X > 500) = e^{-4 \cdot 10^{-4} \cdot 500} = 0,819$$

2. L'affidabilità per $t = 500$ ore di un singolo componente è $R(500) = 0,951$. Allora, la probabilità che il sistema continui a funzionare dopo 500 ore di vita è:
 $R_{3/4}(500) = 4 \cdot R(500)^3 - 3 \cdot R(500)^4 = 0,986$

Osservazione: si veda la prima dispensa per il calcolo dell'affidabilità di un sistema *m out of n*.

3. L'affidabilità per $t = 500$ ore di un singolo componente è $R(500) = 0,951$. Dunque, la probabilità che dopo le 500 ore di vita risultino funzionanti **non più** di due dei quattro componenti è la seguente:

$$R_{0, \dots, 2/4}(500) = \sum_{i=0}^2 \binom{4}{i} R(500)^i \cdot (1 - R(500))^{4-i} = 0,013$$

4. Il tasso di guasto del sistema $h(t)$ alle prime 500 ore di vita si trova come:

$$h(t) = \frac{f_X(t)}{1 - F_X(t)}.$$

Poiché $F_X(t)$ nel caso specifico risulta essere $F_X(t) = 1 - 4 \cdot e^{-3\lambda t} + 3 \cdot e^{-4\lambda t}$, si ricava

$$f_X(t) = 12\lambda \cdot e^{-3\lambda t} - 12\lambda \cdot e^{-4\lambda t}.$$

Quindi, $h(500) = 0,000051$

5. Nel caso in cui all'istante $t = 500$ ore si sia verificato uno ed un solo guasto, il sistema da questo istante in poi sarà la serie di tre componenti. Cambia quindi la $F_X(t)$ cioè $F_X(t) = 1 - e^{-3\lambda t}$.

Il tasso di guasto si determina facilmente considerando la proprietà di assenza di memoria di cui gode l'esponenziale: $h(t) = 3\lambda$.

Osservazione: per ipotesi il sistema funzionerà solo con 3 componenti non guasti.

La distribuzione “normale” (di Gauss).

La distribuzione normale è piuttosto alla base dell’analisi statistica ed ha, invece, un ruolo più limitato nell’analisi probabilistica, non fosse altro che per il fatto che essa, non è esprimibile in forma chiusa (non si possono fare operazioni algebriche) ma solo come funzione integrale della sua densità. Dunque non è manipolabile più di tanto all’interno di formule fondamentali come la distribuzione totale.

Nel linguaggio dell’analisi probabilistica, si può dire che la funzione densità normale è la formalizzazione di un “modo di ripartizione” delle probabilità sullo spazio continuo delle realizzazioni di una variabile aleatoria fondata su due assunzioni:

1. Esiste una “realizzazione tipica”, nello spazio delle possibili realizzazioni, ed è quella che ricorre più frequentemente, valore atteso
2. Esiste un “livello di dispersione ben preciso”, più o meno grande, che dà ragione di scostamenti più o meno grandi dalla realizzazione tipica e, comunque, simmetrici. varianza

La ratio delle precedenti assunzioni si spiega con il fatto che, storicamente, con la densità normale si voleva descrivere un modo di ripartizione degli errori in esperimenti di misura di grandezze fisiche che poggiava sull’idea che l’errore tipico doveva essere nullo; ma ripetendo le misure era inevitabile la presenza di “disturbi di entità limitata”, ma assolutamente casuali e senza un verso preferenziale.

All’ipotesi 1. si può associare, in questa sede, l’interpretazione che la realizzazione tipica sia una sorta di durata vera e propria del fenomeno d’interesse, determinata da fattori endogeni. Mentre, con l’ipotesi 2., si aggiunge il fatto che, in assenza di ulteriori informazioni, si accetta l’idea che anticipi e ritardi (praticamente finiti) sulla durata del fenomeno siano equiprobabili.

Per comodità di trattazione, le ipotesi 1. e 2., saranno per il momento formalizzate imponendo che la forma di densità cercata sia riferita ad una variabile aleatoria, Z , standard, cioè di media nulla, $E[Z] = 0$, e varianza unitaria, $VAR[Z] = 1$. Poi si estenderà la funzione trovata al caso generico e, soprattutto, più naturale in questa sede, di una X con $E[X] = \mu$ e $VAR[X] = \sigma^2$ (si pone al quadrato perché, spesso, si lavora con la radice della varianza, detta deviazione standard e quindi pari a σ).

La media è il punto in cui si centra la Gaussiana. Quindi se media 0 è centrata in ascissa pari a 0. La normale standard/Gaussiana ha deviazione standard 1. La deviazione standard equivale ad un "fattore di distorsione". Nella standard dopo 3 deviazioni, la campana si appiattisce. Quindi se ho media 0 e varianza 16, la non standard con deviazione 4 "muore" a dopo 3 deviazioni, quindi ascissa pari a 12

Proprietà Gaussiana standard

Con le ulteriori condizioni:

per z che tende a infinito
la funzione tende a 0

funzione crescente per $z < 0$ $\frac{d f_Z(z)}{dz} = \begin{cases} > 0 & z < 0 \\ < 0 & z > 0 \end{cases}$ e $f_Z(z \rightarrow \pm \infty) \rightarrow 0$

Più:

condizione per punto di massimo o minimo $\frac{d f_Z(z)}{dz} \Big|_{z=0} = 0$ e $\frac{d^2 f_Z(z)}{dz^2} \Big|_{z=0} < 0$ concava

lasciate all'interpretazione dello studente, si può riconoscere che la più semplice equazione differenziale che le rappresenta tutte è la seguente:

Eq. diff. primo ordine $\rightarrow \frac{d f_Z(z)}{dz} = -f_Z(z) z$; con $f_Z(z) > 0$ per $-\infty < z < +\infty$

Integrando la precedente si ottiene:

$$f_Z(z) = k \exp\{-z^2/2\}, \text{ con } k > 0$$

La costante k si determina imponendo che risulti pari ad uno l'integrale della f esteso a tutto lo spazio delle possibili realizzazioni della Z .

Si pone pari a 1 perché f è una densità ed il suo integrale tra $-\infty$ e $+\infty$ è pari a 1

Dunque: il due è 0 +infinito è perché è simmetrica

$$1 = \int_{-\infty}^{+\infty} k \exp\{-z^2/2\} dz = k \frac{2}{\sqrt{2}} \int_{t=0}^{+\infty} t^{1/2-1} \exp\{-t\} dt \quad (t \triangleq z^2/2)$$

Preso il secondo integrale di Eulero $\int_{t=0}^{+\infty} t^{\alpha-1} \exp\{-t\} dt = \Gamma(\alpha)$
si ottiene il risultato

e poiché riconosciamo la "funzione gamma" nell'ultimo integrale indefinito:

$$\int_{t=0}^{+\infty} t^{1/2-1} \exp\{-t\} dt = \Gamma(1/2) = \sqrt{\pi}$$

concludiamo che la **densità normale standard** è:

rispetto alla non standard al denominatore
sigma non c'è, perché pari a 1 $f_Z(z) = \frac{1}{\sqrt{2\pi}} \exp\{-z^2/2\}, \quad -\infty < z < +\infty$

e da qui la **distribuzione normale standard**, in forma di funzione integrale:

$$F_Z(z) = \int_{-\infty}^{u=z} \frac{1}{\sqrt{2\pi}} \exp\{-u^2/2\} du, \quad -\infty < z < +\infty \quad (\text{rimane funzione integrale})$$

I valori della "normale standard" sono ottenuti per integrazione numerica e sono tabulati. Alcune tabelle riportano la probabilità cumulativa della normale standard, ovvero la probabilità cade tra $-\infty$ and un dato valore $u > 0$, i.e., $P(-\infty \leq z \leq u)$

Osservazione: $F_Z(-z) = 1 - F_Z(z)$ $z \rightarrow 0$ c'è simmetria rispetto all'origine

We shall now prove that the $f_X(x)$ is a probability density function by showing that if

dimostrazione alternativa
dell'integrale precedente

$$I = \int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi}} e^{-(1/2)x^2} dx$$

then $I = 1$. Direct evaluation of the integral is difficult. Hence we square I and write

$$\begin{aligned} I^2 &= \int_{-\infty}^{\infty} \frac{1}{2\pi} e^{-(1/2)x^2} dx \cdot \int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi}} e^{-(1/2)y^2} dy \\ &= \frac{1}{2\pi} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} e^{-(1/2)(x^2+y^2)} dx dy \end{aligned} \quad (6.4.6)$$

Substituting $x = r \cos(\theta)$ and $y = r \sin(\theta)$ in Eq. (6.4.5) we obtain,

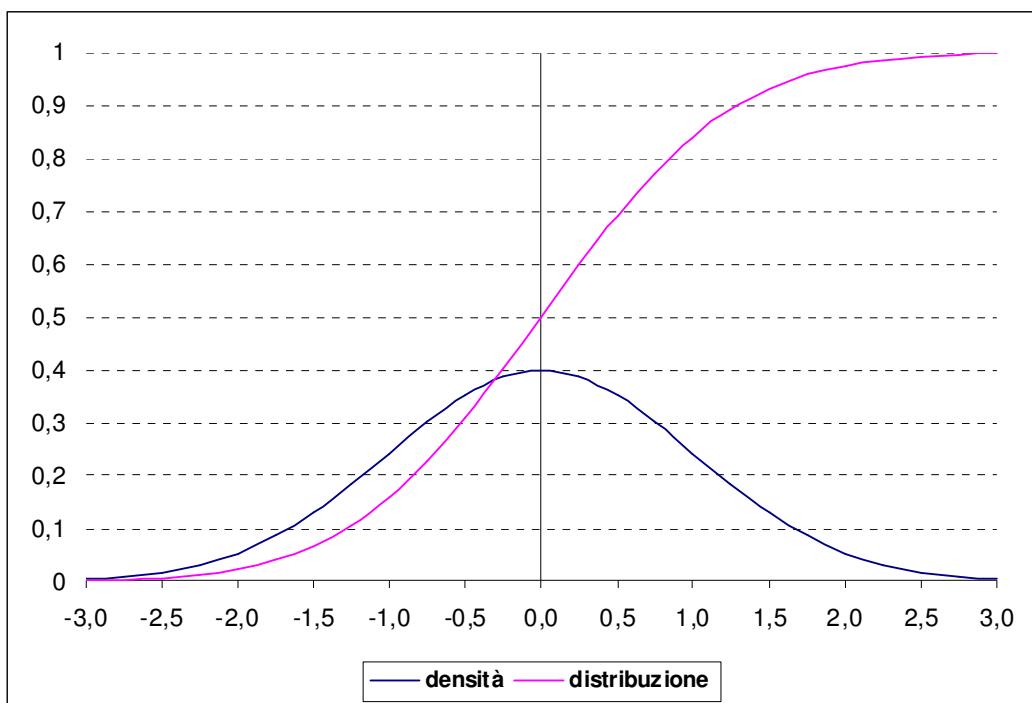
$$I^2 = \frac{1}{2\pi} \int_0^{2\pi} \int_0^{\infty} r e^{-(1/2)r^2} dr d\theta = 1 \quad (6.4.7)$$

and the result is proven.

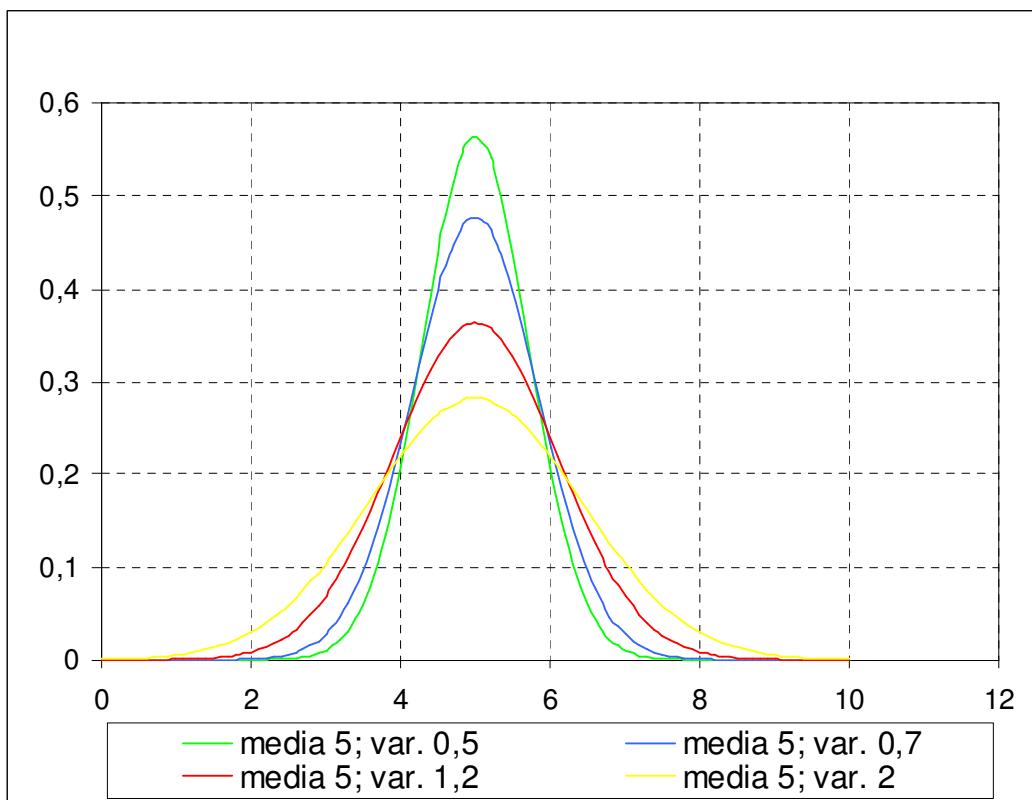
Example 6.4.1 The number of malfunctioning computers is Gaussian-distributed with $\mu = 4$ and $\sigma = 3$. We want to find the probability that the number of bad computers is

Rappresentazioni della normale

è "larga"
3 deviazioni
standard poi
tende a 0



La normale standard.



Esempi di normale NON standard.

DETALGO TECNICO_1

dimostrazione del perché esce sigma al quadrato nella varianza

PER DEFINIZIONE DI VARIANZA:

$$\text{VAR}[x] \triangleq E[(x - E[x])^2]$$

Allora:

$$\text{VAR}[\alpha x] = E[(\alpha x - E[\alpha x])^2]$$

$$= E[(\alpha x - \alpha E[x])^2]$$

$$= E[\alpha^2 \cdot (x - E[x])^2]$$

$$= \alpha^2 E[(x - E[x])^2]$$

$$= \alpha^2 \text{VAR}[x]$$

La distribuzione “normale” (9/10)

Sia $X \stackrel{\text{def}}{=} \sigma Z + \mu$, con $-\infty < \mu < +\infty$ e $\sigma > 0$
variabile aleatoria

X mantiene la forma “normale”, ma è traslata sul valore di μ
scommettiamo che

si dimostra che

$$E[X] = \sigma E[Z] + E[\mu] = \mu$$

Si dimostreranno
in seguito

La larghezza della “campana” è data dal valore di σ

$$\text{VAR}[X] = \sigma^2 \text{VAR}[Z] + \text{VAR}_{\mu} = \sigma^2$$

quelli del grafico precedente

perchè costante

I due punti di flesso corrispondono alle ascisse $\mu \pm \sigma$

creo una relazione tra la gaussiana standard ed una non standard

$$\Pr\{X \leq x\} = \Pr\left\{Z \leq \frac{x - \mu}{\sigma}\right\} \quad \rightarrow \quad z = \frac{x - \mu}{\sigma}$$

questo verrà usato per la non standard

densità non standard sarà:

$$f_X(x) = \frac{1}{\sigma \sqrt{2\pi}} \exp\left\{-\frac{1}{2}\left(\frac{x - \mu}{\sigma}\right)^2\right\}, \quad -\infty < x < +\infty$$

DETTAGLIO TECNICO_2_e_3

utile per la relazione tra standard
e non standard

$$\Pr\{X \leq x\} = \Pr\left\{Z \leq \frac{x - \mu}{\sigma}\right\} \quad \rightarrow \quad z = \frac{x - \mu}{\sigma}$$

$$\begin{aligned} F_X(x) &\hat{=} \Pr\{X \leq x\} = \Pr\{\sigma Z + \mu \leq x\} = \\ &= \Pr\left\{Z \leq \frac{x - \mu}{\sigma}\right\} \hat{=} F_Z\left(\frac{x - \mu}{\sigma}\right) \end{aligned}$$

INOLTRE:

$$\begin{aligned} f_X(x) &\hat{=} \frac{d}{dx} F_X(x) = \frac{d}{dx} F_Z\left(\frac{x - \mu}{\sigma}\right) \\ &= \frac{1}{\sigma} f_Z\left(\frac{x - \mu}{\sigma}\right) \end{aligned}$$

si dimostra
nella lezione
di funzioni di
variabili aleatorie

da qui nasce il
sigma della slide precedente

Risultati fondamentali sulla normale

L'importanza della distribuzione normale è in larga parte dovuta all'esistenza dei due risultati che saranno presentati adesso. Il primo è noto come **teorema di riproducibilità** della distribuzione normale perché stabilisce che la "forma" normale si mantiene rispetto all'operazione di somma (di variabili aleatorie), mentre il secondo è noto come **teorema centrale del limite** perché stabilisce che la legge normale è la forma a cui tende la somma di variabili aleatorie di media e varianza note ed ha, per questo, un ruolo centrale nell'analisi statistica.

Teorema di riproducibilità della normale

Sia X_1, X_2, \dots, X_n una collezione di variabili aleatorie indipendenti e ciascuna distribuita secondo una legge normale di media μ_i e varianza σ_i^2 ($i = 1, \dots, n$). Allora, introdotte le costanti reali a_1, a_2, \dots, a_n , la variabile aleatoria

$$S_n \doteq \sum_{i=1}^n a_i X_i \text{ combinazione lineare}$$

è distribuita con una legge normale di media

$$= \sum_{i=1}^n a_i \mu_i \text{ e varianza } \sigma^2 = \sum_{i=1}^n a_i^2 \sigma_i^2. \text{ combinazioni lineari}$$

Teorema centrale del limite

Sia X_1, X_2, \dots, X_n una collezione di variabili aleatorie indipendenti e ciascuna distribuita secondo una stessa legge, arbitraria e con media μ_i e varianza σ_i^2 anche diverse. Allora la variabile aleatoria:

$$Z_n \text{ tende alla Gaussiana standard} \quad \tilde{Z}_n \doteq \frac{\sum_{i=1}^n X_i - \sum_{i=1}^n \mu_i}{\sqrt{\sum_{i=1}^n \sigma_i^2}} \quad \begin{array}{l} \text{Si sta standardizzando} \\ \text{ad una Gaussiana} \end{array}$$

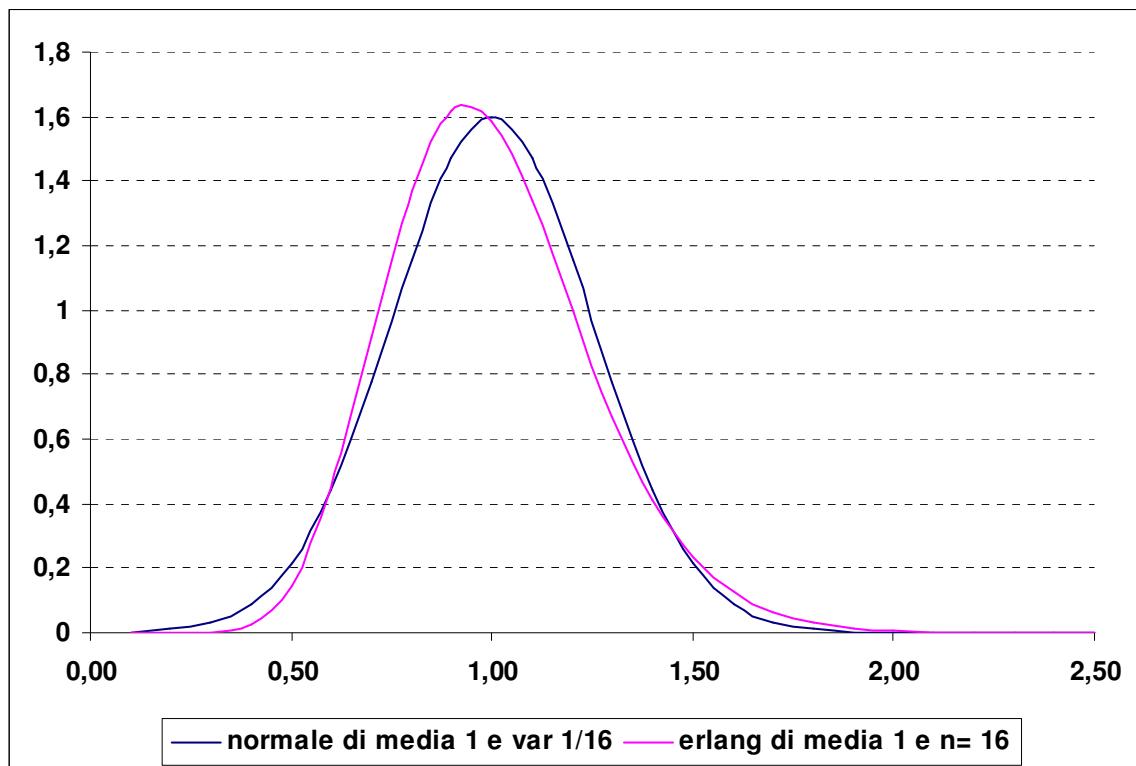
ma la sommatoria delle variabili X_i no

che ha, per costruzione, media nulla e varianza unitaria converge in forma alla normale standard al crescere di n , cioè:

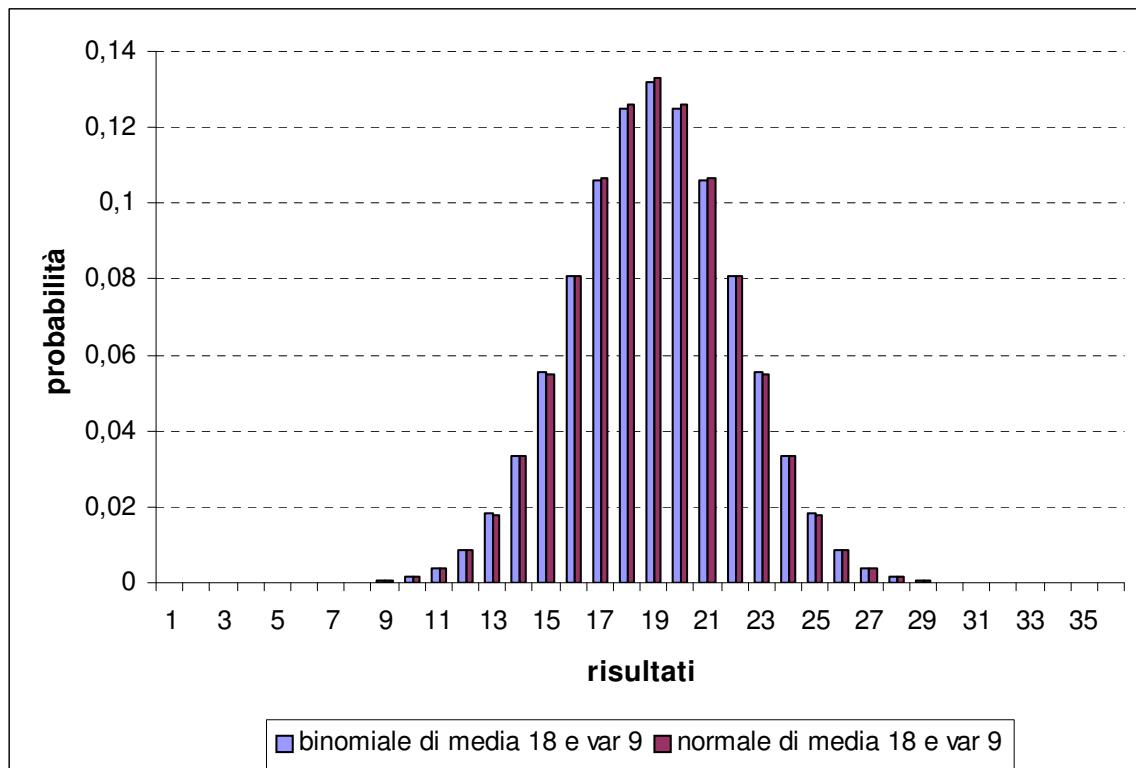
$$\begin{aligned} n \rightarrow \infty \Rightarrow F_{Z_n}(z) &\xrightarrow{\text{distr}} \int_{-\infty}^{u=z} \frac{1}{\sqrt{2\pi}} \exp\{-u^2/2\} du, \quad -\infty < z < +\infty \\ \text{già con } n=30 \\ \text{si ha una normale} \\ \text{La dimostrazione dei due teoremi è tralasciata.} \end{aligned}$$

In pratica prese cose asimmetriche dopo n passaggi ottengono una Gaussiana che è simmetrica

Illustrazione del teorema limite centrale



Erlang è somma di exp come vedremo dopo!



La binomiale è una somma di prove di Bernoulli

Distribuzione congiunta e distribuzione marginale

Il concetto di distribuzione congiunta si pone in due casi:

- 1) quando si voglia definire una successione di variabili aleatorie su di uno stesso spazio dei risultati, al fine di rappresentare la successione delle durate di uno stesso fenomeno che si ripete nel tempo;
- 2) quando si voglia considerare uno spazio dei risultati che sia lo spazio "prodotto" di coppie (o n-ple) di realizzazioni di variabili aleatorie che rappresentano "cose" diverse ma logicamente dipendenti fra di loro.

Esempio tipico del caso 1) :

la successione dei periodi di attività e/o di ozio di una certa risorsa.

Primo esempio per il caso 2):

il numero di utenti già in attesa di essere "processati" da una risorsa e la durata dell'attesa di un utente che si aggiunge a quelli.

Secondo esempio per il caso 2):

i tempi al guasto di due componenti soggetti sia a cause individuali di guasto sia a cause comuni di guasto.

Fissando l'attenzione sul secondo esempio del caso 2), è evidente che lo spazio dei risultati è dato da $\{(t_1, t_2) | 0 \leq t_1 < \infty, 0 \leq t_2 < \infty\}$ e che l'ipotesi di cause comuni di guasto si traduce in una dipendenza logica fra i tempi al guasto (X_1 e X_2) dei due componenti. Dunque, i due eventi " $X_1 \leq t_1$ " e " $X_2 \leq t_2$ ", non sono certamente disgiunti e le rispettive probabilità non possono essere misurate con due funzioni di distribuzione definite separatamente, come modelli primitivi dell'analisi probabilistica.

Si definisce, allora, la funzione di distribuzione congiunta:

$$F_{X_1, X_2}(t_1, t_2) \hat{=} P(X_1 \leq t_1, X_2 \leq t_2), \quad 0 \leq t_1 < \infty, \quad 0 \leq t_2 < \infty$$

come modello primitivo dell'analisi probabilistica degli eventi d'interesse sullo spazio dei risultati elementari: $(t_1, t_2) \equiv (X_1 \leq t_1, X_2 \leq t_2)$ già individuato.

La distribuzione congiunta deve avere tutte le proprietà di una distribuzione e può essere ottenuta a partire dalla definizione preliminare di una funzione densità congiunta, f , alla seguente maniera:

$$F_{X_1 X_2}(t_1, t_2) = \int_{u_1=0}^{t_1} \int_{u_2=0}^{t_2} f_{X_1 X_2}(u_1, u_2) du_2 du_1$$

Esempio

Siano X_1 e X_2 due variabili aleatorie congiuntamente continue con densità di probabilità congiunta data da:

$$f_{X_1, X_2}(x_1, x_2) = \begin{cases} 4e^{-x_1} e^{-2x_2} & x_1 > 0, x_2 > 0 \\ 0 & \text{altrimenti} \end{cases}.$$

Calcolare $P(X_1 > 1, X_2 < 1)$.

$$\begin{aligned} P(X_1 > 1, X_2 < 1) &= \int_{x_2=0}^1 \int_{x_1=1}^{\infty} 4e^{-x_1} e^{-2x_2} dx_1 dx_2 \\ &= \int_{x_2=0}^1 4e^{-2x_2} \left(\int_{x_1=1}^{\infty} e^{-x_1} dx_1 \right) dx_2 \\ &= \int_{x_2=0}^1 4e^{-2x_2} \left(e^{-x_1} \Big|_1^{\infty} \right) dx_2 \\ &= e^{-1} \int_{x_2=0}^1 4e^{-2x_2} dx_2 \\ &= e^{-1} \left(-2e^{-2x_2} \Big|_0^1 \right) = e^{-1} \left(-2e^{-2} + 2 \right) = 2e^{-1} \left(1 - e^{-2} \right) \\ &\approx 0,64 \end{aligned}$$

Gli eventi " $X_1 \leq t_1$ " e " $X_2 \leq t_2$ " sono di particolare interesse perché rappresentano la durata del funzionamento di un singolo componente a prescindere dallo stato in cui si trova l'altro. La probabilità di questi eventi d'interesse è misurata dalle rispettive funzioni di distribuzione marginale:

$$F_{X_i}^{(m)}(t_i) \hat{=} P(X_i \leq t_i) \quad i=1,2$$

Le funzioni di distribuzione marginale si ricavano a partire dalla distribuzione congiunta, come "mostrato" qui di seguito per la X_1 .

Dalle relazioni:

$$(X_1 \leq t_1) \equiv (X_1 \leq t_1 \cap X_2 < \infty) = \lim_{t_2 \rightarrow \infty} (X_1 \leq t_1 \cap X_2 \leq t_2)$$

risulta

$$P(X_1 \leq t_1) = P\left(\lim_{t_2 \rightarrow \infty} (X_1 \leq t_1, X_2 \leq t_2)\right) = \lim_{t_2 \rightarrow \infty} P(X_1 \leq t_1, X_2 \leq t_2)$$

e quindi

$$F_{X_1}^{(m)}(t_1) = \lim_{t_2 \rightarrow \infty} F_{X_1, X_2}(t_1, t_2)$$

Analogamente

$$F_{X_2}^{(m)}(t_2) = \lim_{t_1 \rightarrow \infty} F_{X_1, X_2}(t_1, t_2)$$

Qualora sia data solo la densità congiunta, la distribuzione marginale della X_1 si ottiene dalla seguente:

$$\lim_{t_2 \rightarrow \infty} F_{X_1, X_2}(t_1, t_2) = \int_{u_1=0}^{t_1} \int_{u_2=0}^{\infty} f_{X_1, X_2}(u_1, u_2) du_2 du_1 = F_{X_1}(t_1) \quad (*)$$

e in maniera analoga quella della X_2 .

Infine, osservando la (*) si conviene di definire densità marginale della X_1 la seguente funzione:

$$f_{X_1}(u_1) \hat{=} \int_{u_2=0}^{\infty} f_{X_1, X_2}(u_1, u_2) du_2 < \infty$$

che deve avere tutte le proprietà di una funzione densità.

In maniera del tutto analoga, lo studente può definire la densità marginale della X_2 .

La distribuzione congiunta in Affidabilità

Siano X_1 e X_2 variabili aleatorie che misurano, rispettivamente, il tempo al guasto del primo e del secondo di due componenti particolari, all'interno di un generico sistema elettronico. Allora, la distribuzione congiunta:

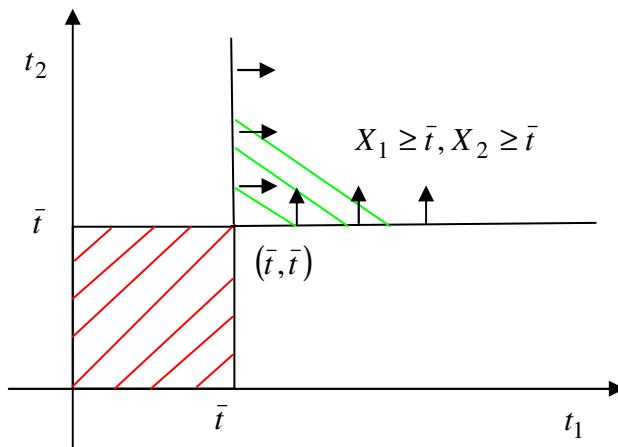
$$F_{X_1, X_2}(t, t) \hat{=} P(X_1 \leq t \cap X_2 \leq t) \quad 0 \leq t < \infty$$

può essere usata per esprimere la probabilità che entrambi i componenti non sopravvivano oltre l'istante t , nell'ipotesi che agiscano sia cause di guasto individuale sia cause di guasto comune. Più in generale, la distribuzione congiunta può essere adottata come modello primitivo quando non è possibile caratterizzare separatamente il tempo al guasto di uno qualunque dei due componenti, indipendentemente dall'altro.

In tale circostanza, se si vuole definire “l'affidabilità congiunta” dei due componenti, come probabilità che entrambi superino il generico istante t in condizioni di buon funzionamento, si può adottare la funzione seguente:

$$R(t, t) \hat{=} P(X_1 > t \cap X_2 > t)$$

Osservando la figura seguente si evita l'errore: $R(\bar{t}, \bar{t}) = 1 - F_{X_1, X_2}(\bar{t}, \bar{t})$ (falso!)



e si può riconoscere invece la seguente:

$$P(X_1 \cap X_2 > \bar{t}) \hat{=} \int_{u_1=\bar{t}}^{\infty} \int_{u_2=\bar{t}}^{\infty} f_{X_1, X_2}(u_1, u_2) du_1 du_2$$

Variabili aleatorie indipendenti

Se e solo se risulta:

$$F_{X_1, X_2}(t_1, t_2) = F_{X_1}^{(m)}(t_1) F_{X_2}^{(m)}(t_2), \quad 0 \leq t_1 < \infty, \quad 0 \leq t_2 < \infty$$

ovvero

$$f_{X_1, X_2}(t_1, t_2) \doteq f_{X_1}^{(m)}(t_1) f_{X_2}^{(m)}(t_2),$$

allora le variabili aleatorie X_1 e X_2 sono dette indipendenti.

Particularizzazione della ripartizione congiunta e marginale al caso discreto

Nel caso in cui si abbia a che fare con variabili aleatorie discrete, ad esempio N e K indicanti, rispettivamente, il numero di oggetti presenti in due buffer comunicanti secondo uno schema produttore-consamatore sono utili le seguenti:

$$P_{N,K}(n,k) \doteq \text{Prob}\{N = n, K = k\}, \quad n = 0, 1, 2, \dots ; \quad k = 0, 1, 2, \dots ;$$

$$P_N^{(m)}(n) \doteq \sum_{k=0}^{\infty} P_{N,K}(n,k) \quad C_N^{(m)}(l) = \sum_{n=0}^l P_N^{(m)}(n)$$

lasciate all'interpretazione dello studente, per esercizio.

La distribuzione condizionata

Date due variabili aleatorie X e Y , con Y continua, la funzione di distribuzione condizionata $F_{Y/X}(y/x)$ viene introdotta per misurare la probabilità dell'evento " $Y \leq y / X = x$ ". Per proporre una definizione formale che si riferisca al caso in cui la X è una variabile aleatoria continua con funzione di distribuzione ottenuta integrando la densità, si può pensare all' evento condizionante come ad un evento limite :

$$"Y \leq y / X = x" \hat{=} \lim_{\Delta x \rightarrow 0} "Y \leq y / x \leq X \leq x + \Delta x"$$

e assumere che valga la seguente:

$$P(\lim_{\Delta x \rightarrow 0} "Y \leq y / x \leq X \leq x + \Delta x") = \lim_{\Delta x \rightarrow 0} P(Y \leq y / x \leq X \leq x + \Delta x)$$

Cosicché :

$$F_{Y/X}(y/x) \hat{=} \lim_{\Delta x \rightarrow 0} P(Y \leq y / x \leq X \leq x + \Delta x)$$

Con la definizione proposta, la distribuzione condizionata è posta in relazione ad un evento condizionante la cui probabilità è misurabile con gli strumenti a nostra disposizione e si può ricavare la seguente formula di calcolo:

$$F_{Y|X}(y|x) = \int_{v=0}^y \frac{f_{Y,X}(v,x)}{f_X^{(m)}(x)} dv$$

Da cui è immediato riconoscere la funzione densità condizionata :

$$f_{Y/X}(v/x) \hat{=} \frac{f_{Y,X}(x,v)}{f_X^{(m)}(x)}$$

Derivazione della formula di calcolo

Prima di ricavarla, è bene far vedere che la probabilità dell'evento condizionante è misurabile con la densità marginale:

$$\begin{aligned} P(x \leq X \leq x + \Delta x) &= \lim_{y \rightarrow \infty} P(x \leq X \leq x + \Delta x \cap Y \leq y) \\ &= \int_{u=x}^{x+\Delta x} \int_{v=0}^{\infty} f_{Y,X}(v,u) dv du = f_X^{(m)}(x) \Delta x \end{aligned}$$

A questo punto:

$$\begin{aligned} F_{Y/X}(y/x) &\triangleq \lim_{\Delta x \rightarrow 0} P(Y \leq y / x \leq X \leq x + \Delta x) \\ &= \lim_{\Delta x \rightarrow 0} \frac{P(Y \leq y \cap x \leq X \leq x + \Delta x)}{\lim_{y \rightarrow \infty} P(x \leq X \leq x + \Delta x \cap Y \leq \infty)} \\ &= \lim_{\Delta x \rightarrow 0} \frac{\int_{v=0}^y \int_{u=x}^{x+\Delta x} f_{Y,X}(v,u) du dv}{\Delta x f_X^{(m)}(x)} \\ &= \lim_{\Delta x \rightarrow 0} \frac{\int_{v=0}^y f_{Y,X}(v, \bar{u}) \Delta x \, dv}{\Delta x f_X^{(m)}(x)} \quad \text{ove } x \leq \bar{u} \leq x + \Delta x \\ &= \int_{v=0}^y \frac{f_{Y,X}(v,x)}{f_X^{(m)}(x)} \, dv \end{aligned}$$

Formula della distribuzione totale

La formula della distribuzione totale mette in relazione la funzione di distribuzione marginale $F_Y^{(m)}(y)$ con la funzione di distribuzione condizionata $F_{Y|X}(y/x)$, purché si conosca pure la densità marginale della X . È una formula alla quale si ricorre spesso, perché è quasi sempre più facile conoscere direttamente la distribuzione condizionata che non la marginale. Da un punto di vista teorico, non è altro che l'estensione della formula della probabilità totale alle variabili aleatorie.

Per ricavarla, si può partire dalla relazione:
$$F_Y^{(m)}(y) = \lim_{x \rightarrow \infty} F_{Y,X}(y,x)$$

Ricordando che:

$$F_{Y,X}(y,x) \triangleq \int_{v=0}^y \int_{u=0}^x f_{Y,X}(v,u) du dv$$

si ha:

$$\begin{aligned} F_Y^{(m)}(y) &= \int_{v=0}^y \int_{x=0}^{\infty} f_{Y,X}(v,x) dx dv = \int_{v=0}^y \int_{x=0}^{\infty} f_{Y/X}(v/x) f_X^{(m)}(x) dx dv \\ &= \int_{x=0}^{\infty} \int_{v=0}^y f_{Y/X}(v/x) dv f_X^{(m)}(x) dx \end{aligned}$$

e da qui la formula finale:
$$F_Y^{(m)}(y) = \int_{x=0}^{\infty} F_{Y/X}(y/x) f_X^{(m)}(x) dx$$

A parole: la probabilità del risultato $Y \leq y$ è pari alla somma - su tutti i risultati condizionanti $0 \leq x < \infty$ - dei prodotti tra la probabilità di ciascun risultato condizionato $Y \leq y | X \in (x, x+dx)$ e la probabilità del rispettivo condizionante $X \in (x, x+dx)$.

Infine, per derivazione della formula finale, si ottiene la formula della densità totale:

$$f_Y^{(m)}(y) = \int_{x=0}^{\infty} f_{Y/X}(y/x) f_X^{(m)}(x) dx$$

Quando la variabile condizionante X ha realizzazioni non negative, ma nel discreto, allora le formule di distribuzione totale e di densità totale diventano le seguenti:

$$F_Y^{(m)}(y) = \sum_{x=0}^{\infty} F_Y(y|x) \cdot P_X^{(m)}(x) \quad f_Y^{(m)}(y) = \sum_{x=0}^{\infty} f_Y(y|x) \cdot P_X^{(m)}(x).$$

Un'applicazione con variabile condizionante discreta

Lungo la banchina di un porto ha appena ormeggiato una nave portacontainer. Per le operazioni di scarico e carico è previsto l'impiego, in parallelo, di 2 gru identiche. Una serie di fenomeni possono intervenire per allungare i tempi di lavoro di una gru su una nave: un operatore di esperienza ha valutato in 0,2 la probabilità che questo accada e ha pure concluso che, in tale evenienza, la durata media delle operazioni di scarico e carico passa da una media di 4 ad una media di 6 ore. Qual è il modello di calcolo della durata media delle operazioni? E della funzione di distribuzione della stessa durata?

Sia Y la durata delle operazioni di carico/scarico ed X la variabile aleatoria (discreta) condizionante. X ha 2 realizzazioni:

$$X = \begin{cases} 0 & \text{non si verificano ritardi} \\ 1 & \text{si verificano ritardi} \end{cases}$$

con $P(X = 0) = 0,8$ e $P(X = 1) = 0,2$, le cui funzioni di distribuzione sono rispettivamente $F_{Y|X=0} = 1 - e^{-\frac{1}{4}t}$ e $F_{Y|X=1} = 1 - e^{-\frac{1}{6}t}$.

È ora possibile adottare la formula della distribuzione totale quale modello di calcolo della distribuzione della durata delle operazioni di carico/scarico:

$$\begin{aligned} F_Y^{(m)}(t) &= F_{Y|X=0} \cdot P(X = 0) + F_{Y|X=1} \cdot P(X = 1) \\ &= \left(1 - e^{-\frac{1}{4}t}\right) \cdot 0,8 + \left(1 - e^{-\frac{1}{6}t}\right) \cdot 0,2 \\ &= 0,8 - 0,8e^{-\frac{1}{4}t} + 0,2 - 0,2e^{-\frac{1}{6}t} \\ &= 1 - \left(0,8e^{-\frac{1}{4}t} + 0,2e^{-\frac{1}{6}t}\right). \end{aligned}$$

Nota la formula della distribuzione totale, facendone il valore atteso si ottiene la durata media delle operazioni:

$$E[Y] = \int_{t=0}^{\infty} (1 - F_Y^{(m)}(t)) dt$$

$$\begin{aligned}
&= \int_{t=0}^{\infty} \left(1 - 1 + 0,8e^{-\frac{1}{4}t} - 0,2e^{-\frac{1}{6}t} \right) dt \\
&= \left(0,8e^{-\frac{1}{4}t} - 0,2e^{-\frac{1}{6}t} \right) \Big|_0^{\infty} = \left(-0,8 \cancel{\Big/} \cancel{\Big/} \frac{1}{4} e^{-\frac{1}{4}t} - 0,2 \cancel{\Big/} \cancel{\Big/} \frac{1}{6} e^{-\frac{1}{6}t} \right) \Big|_0^{\infty} \\
&= 3,2 + 1,2 = 4,4h
\end{aligned}$$

Un'applicazione con variabile condizionante continua

Si consideri il problema di un utente che arriva di fronte a due risorse, per lui equivalenti ma entrambe occupate. Dovendo scegliere in maniera irrevocabile la risorsa di fronte alla quale accodarsi e nell'ipotesi che le due risorse abbiano due velocità di lavoro diverse (v_1 e v_2), l'utente potrebbe decidere in base alla: *Probabilità ("una risorsa completa il lavoro in atto prima dell'altra")*.

Detta X_i "la durata del lavoro residuo della risorsa i , $i=1,2$ " , si calcola:

$$\begin{aligned} \Pr\{X_2 > X_1\} &= \int_{t=0}^{\infty} \lim_{\Delta t \rightarrow 0} \{P(X_2 > t | t < X_1 \leq t + dt) P(t < X_1 \leq t + dt)\} \\ &= \int_{t=0}^{\infty} [1 - F_{X_2|X_1}(t)] f_{X_1}^{(m)}(t) dt \end{aligned}$$

Aggiungendo l'ipotesi che le due risorse lavorano in modo indipendente risulterebbe:

$$F_{X_2|X_1}(t) = F_{X_2}(t) \quad f_{X_1}^{(m)} = \frac{f_{X_1, X_2}}{f_{X_2/X_1}} = f_{X_1}$$

$$\text{e quindi } \Pr\{X_2 > X_1\} = \int_0^{\infty} [1 - F_{X_2}(t)] f_{X_1}(t) dt$$

Da qui, adottando la legge esponenziale per le variabili in gioco si può ottenere la seguente formula finale:

$$\Pr\{X_1 < X_2\} = \int_0^{\infty} e^{-v_2 t} v_1 e^{-v_1 t} dt = \frac{1}{v_1 + v_2}$$

In modo del tutto analogo

$$\Pr\{X_2 < X_1\} = \frac{v_2}{v_1 + v_2}$$

e, in conclusione, sarebbe giustificata dalle ipotesi introdotte la scelta "naturale" di accodarsi di fronte alla risorsa più veloce.

Distribuzione della somma di variabili aleatorie indipendenti

Si vuole determinare la funzione di distribuzione della variabile aleatoria definita come somma di altre due:

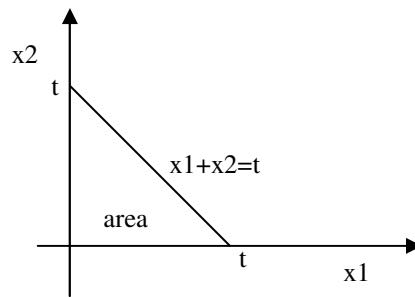
$$X \triangleq X_1 + X_2 ,$$

sullo spazio dei risultati dato dalle coppie reali non negative: $\{(x_1, x_2); x_1 \geq 0, x_2 \geq 0\}$

Considerato che

$$F_X(t) \triangleq \Pr\{X_1 + X_2 \leq t\}$$

occorre integrare la densità congiunta, $f_{X_1 X_2}$, sull'area mostrata in figura:



Ricordando le regole d'integrazione, si scrive subito:

$$F_X(t) = \int_{x_2=0}^t \int_{x_1=0}^{t-x_2} f_{X_1, X_2}(x_1, x_2) dx_1 dx_2 \equiv \int_{x_1=0}^t \int_{x_2=0}^{t-x_1} f_{X_1, X_2}(x_1, x_2) dx_2 dx_1$$

Da qui, se X_1 e X_2 sono variabili aleatorie indipendenti si ha:

$$\begin{aligned} \int_{u_1=0}^t \int_{u_2=0}^{t-u_1} f_{X_1, X_2}(u_1, u_2) du_1 du_2 &= \int_{u_1=0}^t \left(\int_{u_2=0}^{t-u_1} f_{X_2}(u_2) du_2 \right) f_{X_1}(u_1) du_1 \\ &= \int_{x_1=0}^t F_{X_2}(t-x_1) f_{X_1}(x_1) dx_1 \end{aligned}$$

Ovvero

$$= \int_{x_2=0}^t F_{X_1}(t-x_2) f_{X_2}(x_2) dx_2$$

ovvero la convoluzione delle funzioni densità dell'una con la distribuzione dell'altra.

Esempio

Lungo la banchina di un porto ha appena ormeggiato una nave portacontainer. Per le operazioni di scarico e carico container è previsto l'impiego di una gru. L'operatore che manovra detta gru è sottoposto ad un turno di lavoro di 6 ore, che è iniziato proprio con l'avvio delle operazioni di scarico e carico su questa nave il cui tempo medio è stimato in 4 ore. La gru dovrà poi traslare su un punto di ormeggio adiacente per scaricare una seconda nave, con un tempo medio di scarico che, in questo secondo caso, è stimato pari a 2 ore. Trascurando il tempo necessario alla traslazione, proporre un modello per valutare la probabilità che il gruista riesca a completare anche il lavoro sulla seconda nave, entro il turno di lavoro attuale.

Sia

Y_1 la durata delle operazioni di scarico/carico container da effettuare sulla prima nave e Y_2 la durata delle operazioni di scarico/carico container da effettuare sulla seconda nave, con funzioni di distribuzione rispettivamente $F_{Y_1}(t) = 1 - e^{-\frac{1}{4}t}$ con $\mu_1 = \frac{1}{4}$ ore $^{-1}$ e

$$F_{Y_2}(t) = 1 - e^{-\frac{1}{2}t} \text{ con } \mu_2 = \frac{1}{2} \text{ ore}^{-1}.$$

Indicando con $Y = Y_1 + Y_2$ la durata complessiva delle operazioni sulle due navi, allora

$$\begin{aligned} P(Y \leq 6) &= P(Y_1 + Y_2 \leq 6) = F_Y(6) \\ &= \int_{y_1=0}^6 F_{Y_2}(t - y_1) f_{Y_1}(y_1) dy_1 = \int_{y_1=0}^6 (1 - e^{-\frac{1}{2}(t-y_1)}) \cdot \frac{1}{4} e^{-\frac{1}{4}y_1} dy_1 \\ &= \int_{y_1=0}^6 \frac{1}{4} e^{-\frac{1}{4}y_1} dy_1 - \int_{y_1=0}^6 \frac{1}{4} e^{-\frac{1}{4}y_1} e^{-\frac{1}{2}(t-y_1)} dy_1 \\ &= -e^{-\frac{1}{4}y_1} \Big|_0^6 - \frac{1}{4} \bar{e}^{-\frac{1}{2}t} \int_{y_1=0}^6 \bar{e}^{(\frac{1}{4}-\frac{1}{2})y_1} dy_1 \\ &= -e^{-\frac{1}{4}y_1} \Big|_0^6 + \frac{1}{4} \bar{e}^{-\frac{1}{2}t} \bar{e}^{(\frac{1}{4}-\frac{1}{2})y_1} \Big|_0^6 \end{aligned}$$

Convoluzione di esponenziali identiche

Si supponga di avere la collezione { X_1, X_2, \dots, X_n } di variabili aleatorie indipendenti ed identicamente distribuite con legge esponenziale, $F_{X_i}(t) = 1 - e^{-\lambda t}$ $i = 1, 2, \dots, n$. Allora:

$$f_X(t) = \frac{(\lambda t)^{n-1}}{(n-1)!} \lambda e^{-\lambda t}, \quad \text{dove } X \doteq X_1 + X_2 + \dots + X_n$$

Infatti, ragionando per induzione e ponendo:

$$Y \doteq X_1 + X_2 + \dots + X_{n-1} \Rightarrow X = Y + X_n$$

si ipotizza la seguente: $f_Y(t) = \frac{(\lambda t)^{n-2}}{(n-2)!} \lambda e^{-\lambda t}$

e si verifica che:
$$\begin{aligned} f_X(t) &= \int_{u=0}^t f_{X_n}(t-u) f_Y(u) du \quad (\text{perché } X \doteq X_1 + X_2 + \dots + X_n) \\ &= \int_{u=0}^t \lambda e^{-\lambda(t-u)} e^{-\lambda u} \frac{(\lambda u)^{n-2}}{(n-2)!} d(\lambda u) = \lambda e^{-\lambda t} \frac{(\lambda t)^{n-1}}{(n-1)!} \end{aligned}$$

La funzione di distribuzione si ricava integrando la densità:

$$\begin{aligned} F_X(t) &= \int_0^t e^{-\lambda u} \frac{(\lambda u)^{n-1}}{(n-1)!} d(\lambda u) \\ &= e^{-\lambda t} \frac{(\lambda t)^n}{n!} + \int_0^t e^{-\lambda u} \frac{(\lambda u)^n}{n!} d(\lambda u) = e^{-\lambda t} \frac{(\lambda t)^n}{n!} + \int_0^t e^{-\lambda u} d\left(\frac{(\lambda u)^{n+1}}{(n+1)!}\right) \\ &= e^{-\lambda t} \frac{(\lambda t)^n}{n!} + e^{-\lambda t} \frac{(\lambda t)^{n+1}}{(n+1)!} + \int_0^t e^{-\lambda u} d\left(\frac{(\lambda u)^{n+2}}{(n+2)!}\right) = \dots \\ ... &= e^{-\lambda t} \sum_{i=n}^{\infty} \frac{(\lambda t)^i}{i!} \end{aligned}$$

e poiché

$$\sum_{i=0}^{\infty} \frac{(\lambda t)^i}{i!} = e^{\lambda t} \quad \text{si ottiene: } F_X(t) = 1 - e^{-\lambda t} \sum_{i=0}^{n-1} \frac{(\lambda t)^i}{i!}$$

che è conosciuta come distribuzione di Erlang (di ordine n).

Illustrazione del Modello di Erlang (1)

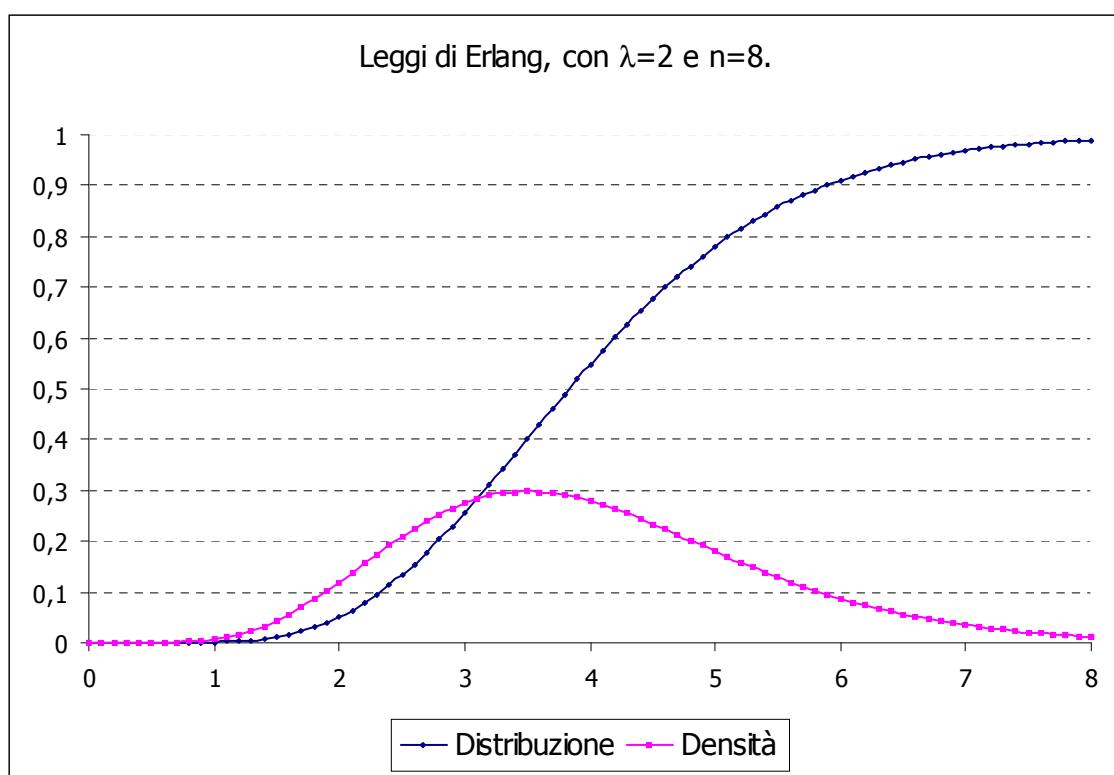
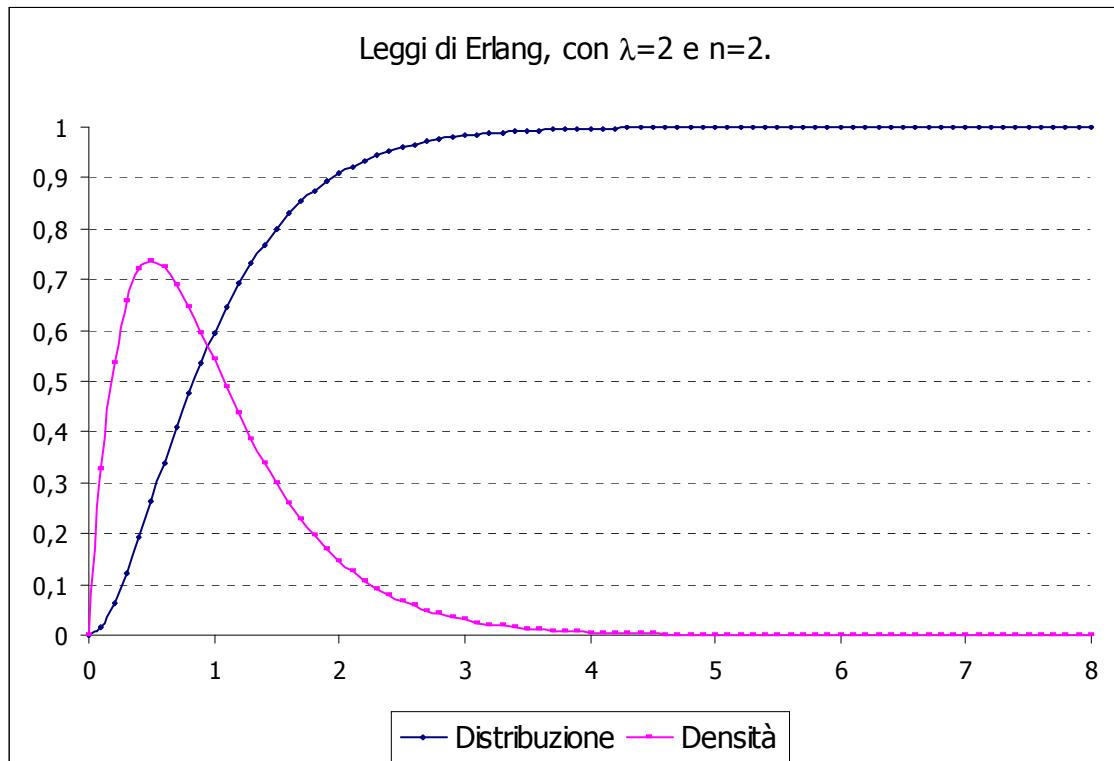
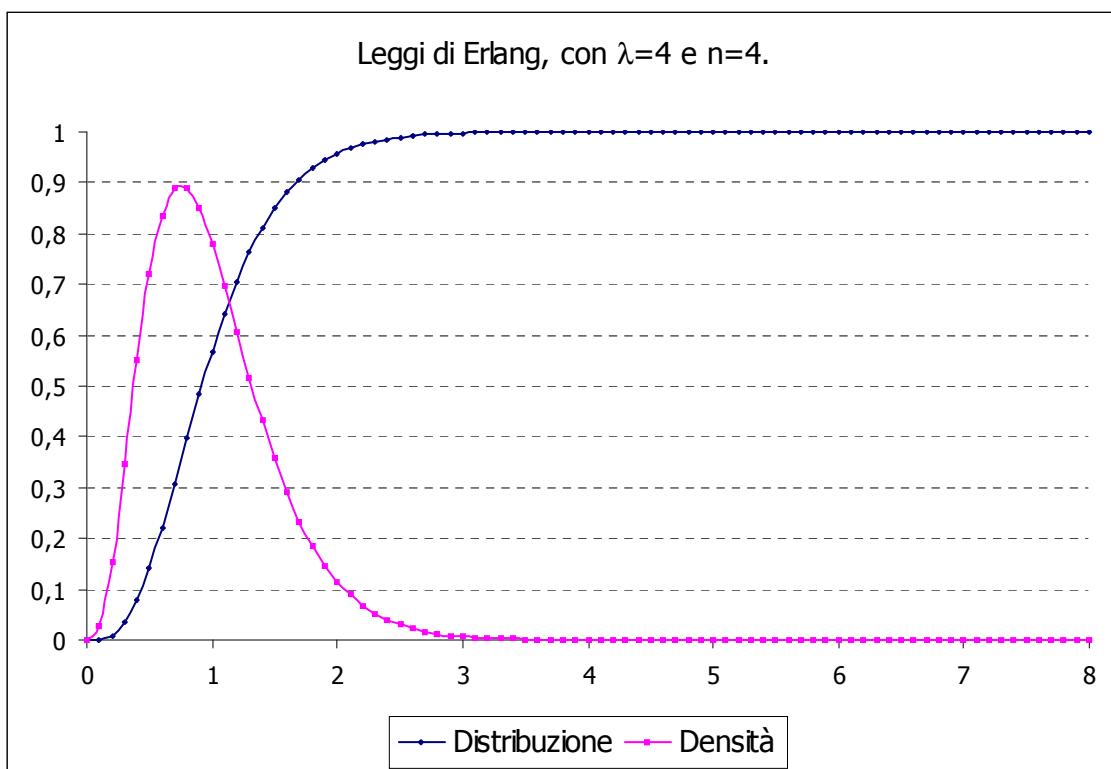
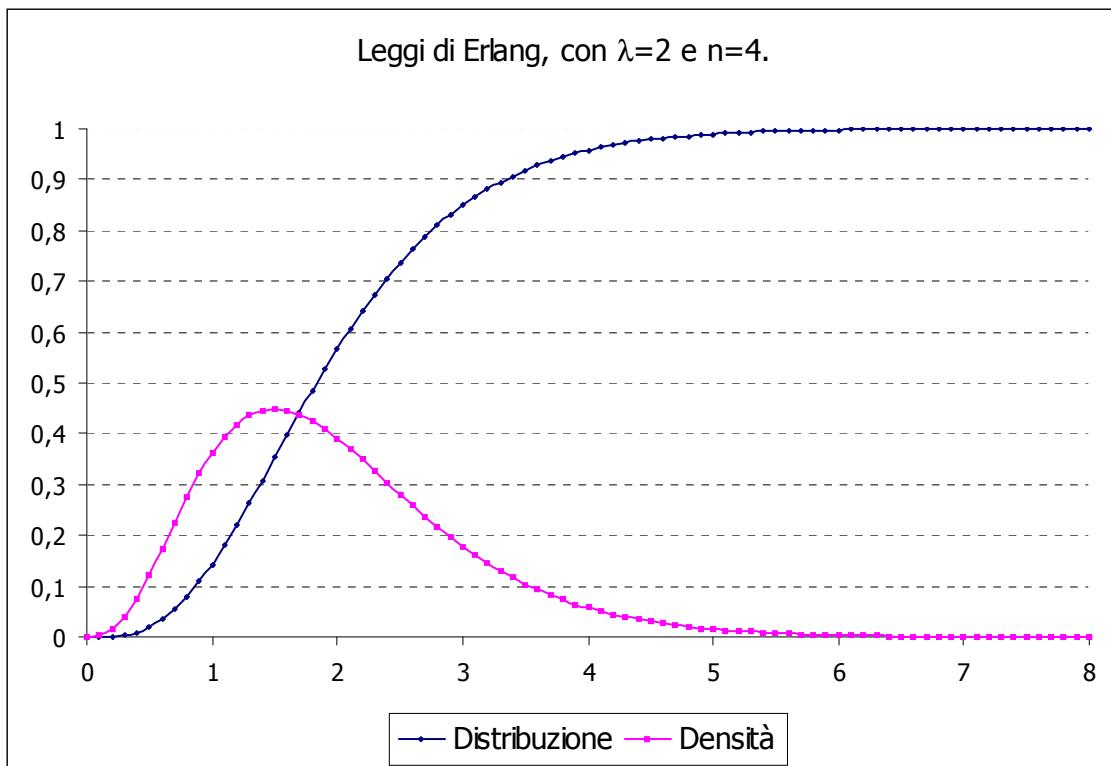
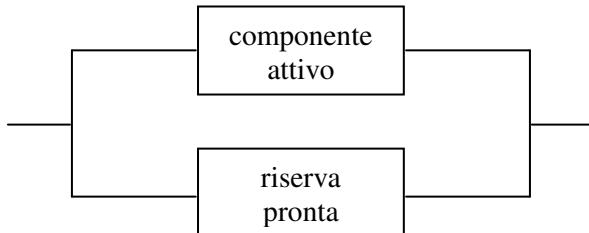


Illustrazione del Modello di Erlang (2)



Affidabilità dei sistemi con riserve pronte

Il più semplice sistema costituito da un componente “attivo” e da una riserva “pronta” è illustrato in figura.



Il componente è attivo nel senso che è messo in funzione fin dall'inizio, mentre la riserva è detta pronta per intendere che è pronta a sostituire il componente attivo, non appena questo subisca un guasto. Nei sistemi elettrici, la riserva è un componente alimentato anche quando è in stato di “pronto” e, dunque, il commutatore raffigurato provvederà a porla in stato di attività.

In prima approssimazione, si assume che la commutazione sia istantanea e che riesca con certezza (commutazione perfetta).

Assumendo, ancora, che la riserva non possa subire guasti quando permane pronta, detta X la variabile “tempo al guasto del sistema”, risulterà $X = X_1 + X_2$ dove X_1 è il tempo di vita del componente attivo e X_2 quello della riserva.

Aggiungendo l’ipotesi di indipendenza si ha:

$$R(t) = 1 - F_X(t) \quad \text{dove } F_X(t) = \int_{x_1=0}^t [1 - e^{-\lambda_2(t-x_1)}] \lambda_1 e^{-\lambda_2 x_1} dx_1 .$$

Nell’ipotesi che il componente attivo e la riserva siano entrambi modellabili con una legge esponenziale, sia pure di parametro diverso (λ_1 e λ_2) si ha:

$$F_X(t) = \int_{x_1=0}^t [1 - e^{-\lambda_2(t-x_1)}] \lambda_1 e^{-\lambda_1 x_1} dx_1 = 1 - \left(\frac{\lambda_2}{\lambda_2 - \lambda_1} e^{-\lambda_1 t} + \frac{\lambda_1}{\lambda_1 - \lambda_2} e^{-\lambda_2 t} \right)$$

Se, come caso particolare, si assumono gli stessi tassi di guasto per il componente attivo e per la riserva ($\lambda_1 = \lambda_2 = \lambda$) allora la F del tempo di vita del sistema risulta la legge di Erlang di ordine 2:

$$F_X(t) = 1 - e^{-\lambda t} \sum_{i=0}^{2-1} \frac{(\lambda t)^i}{i!} = 1 - e^{-\lambda t} (1 + \lambda t)$$

Generalizzazione: da 1 componente di riserva a n-1

Si consideri un sistema composto da "n" componenti, di cui uno è il componente attivo e gli altri "n-1" sono di *riserva*: in caso di malfunzionamento del componente attivo, un commutatore permette di attivare la riserva, al posto del componente attivo.

Sotto le seguenti ipotesi:

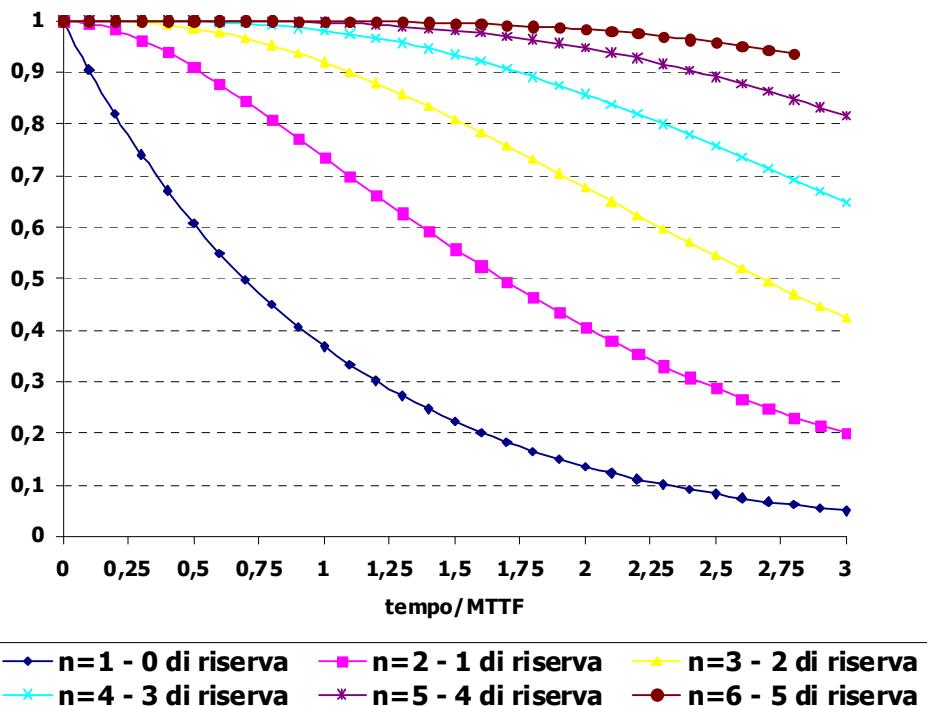
1. *la commutazione riesce sicuramente;*
2. *la commutazione si realizza in un tempo trascurabile ;*
3. *i componenti sono identici e non riparabili;*
4. *la legge comune di guasto è esponenziale di parametro λ ;*

il "tempo al guasto" del sistema segue la legge di Erlang di ordine n e di parametro λ ,

Dunque:

$$R(t) = e^{-\lambda t} \sum_{i=0}^{n-1} \frac{(\lambda t)^i}{i!}$$

**Affidabilità di sistema con un componente attivo e n-1 di riserva
(modello di Erlang)**



Applicazione: Commutazione imperfetta ed Affidabilità

Si riprenda in considerazione il sistema con un componente attivo e una riserva, e si rimuova l'ipotesi che la commutazione sulla riserva riesca sempre. Si assuma, invece, che la commutazione riesce con probabilità "c" e non riesce con probabilità "1-c".

Introducendo la v.a. X , definita come segue:

$$X = \begin{cases} 1 & \text{comutazione riesce} \\ 0 & \text{comutazione fallisce} \end{cases}$$

e con il seguente modello bernoulliano:

$$P_X(1) = c, \quad P_X(0) = 1 - c,$$

si osserva che la variabile aleatoria $Y = \text{"tempo di vita del sistema"}$ (non riparabile) è condizionata dalla X .

In particolare, quando la commutazione non riesce ($X=0$) si ha:

$$F_{Y|X=0}(t) = 1 - e^{-\lambda t};$$

mentre se $X=1$ si ha:

$$F_{Y|X=1}(t) = 1 - e^{-\lambda t} \sum_{i=0}^{2-1} \frac{(\lambda t)^i}{i!} = 1 - e^{-\lambda t} (1 + \lambda t).$$

Allora si applica la formula della distribuzione totale

$$F_Y^{(m)}(t) = \sum_{x=0}^1 F_{Y|X=x}(t) \cdot P_X(x) = F_{Y|X=1}(t) \cdot P_X(1) + F_{Y|X=0}(t) \cdot P_X(0)$$

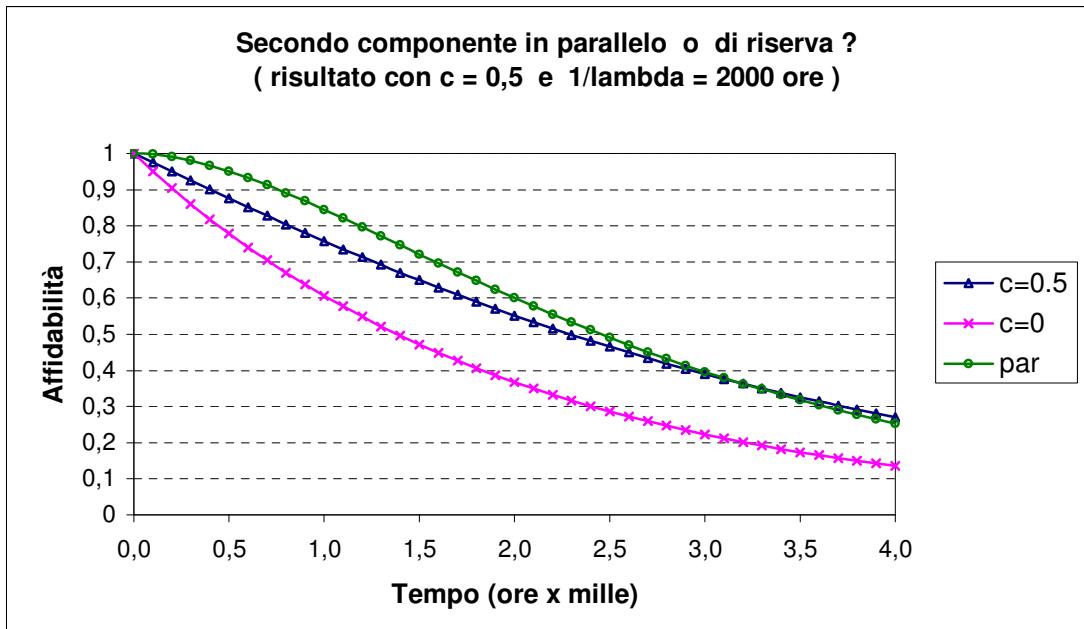
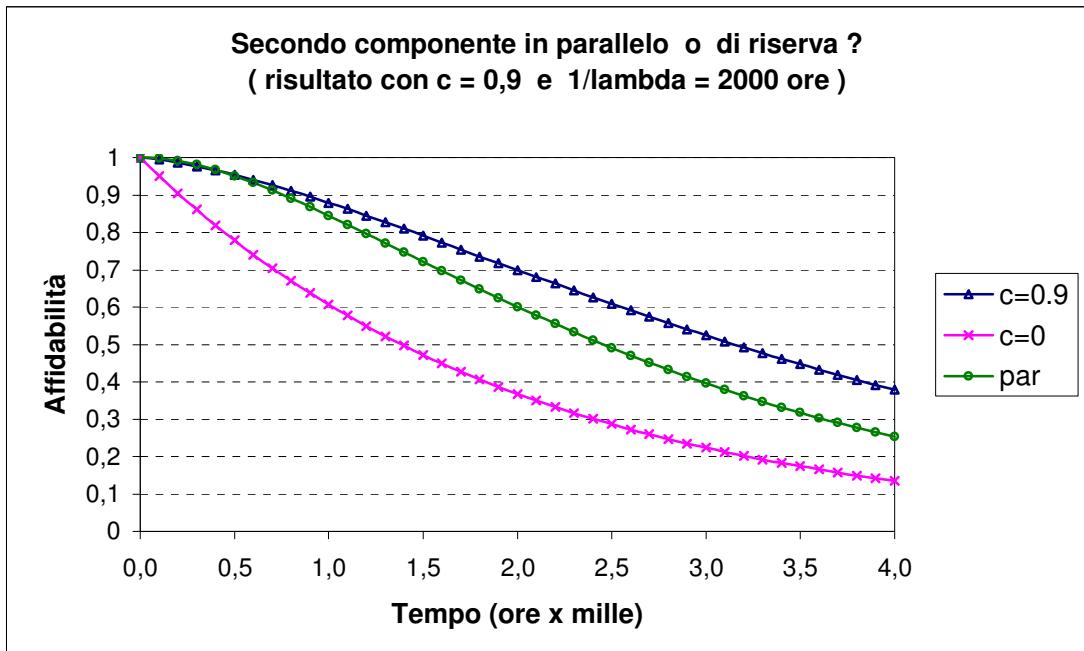
e si ottiene:

$$F_Y^{(m)}(t) = (1 - e^{-\lambda t} - \lambda t e^{-\lambda t}) \cdot c + (1 - e^{-\lambda t}) \cdot (1 - c) = 1 - (1 + c\lambda t) e^{-\lambda t}.$$

Dunque la formula dell'affidabilità del sistema è la seguente:

$$R(t) = 1 - F_Y^{(m)}(t) = (1 + c\lambda t) e^{-\lambda t}, \quad t \geq 0.$$

Illustrazione del “peso” della commutazione



Media e varianza

In precedenza è stata definito il valore atteso della variabile aleatoria X, ed è stato indicato con $E[X]$.

Formule utili

Per variabili aleatorie continue sono valide le seguenti relazioni, che possono essere riscritte anche per variabili discrete:

$$E[X + Y] = E[X] + E[Y] \quad (\text{proprietà di additività})$$

$$E[cX + d] = cE[X] + d \quad (c \text{ e } d \text{ costanti arbitrarie})$$

e aggiungendo l'ipotesi di indipendenza

$$E[XY] = E[X]E[Y] \quad (\text{proprietà di moltiplicatività})$$

Le precedenti rimangono valide passando da due a "n" variabili.

La **varianza** di una variabile aleatoria X è definita come:

$$VAR[X] \equiv E[(X - E[X])^2]$$

La precedente espressione risulta pari a:

$$VAR[X] = E[X^2 - 2XE[X] + E[X]^2] = \dots = E[X^2] - E[X]^2$$

Con riferimento ad una coppia di variabili aleatorie continue, X e Y, allo spazio delle realizzazioni congiunte $\{(x, y) \in \mathbb{R}^2 \mid 0 \leq x < \infty, 0 \leq y < \infty\}$ e ad una generica funzione delle due, $g(X, Y)$, si definisce **valore atteso o media** della g e si indica con $E[g(X, Y)]$ il valore finito (quando esiste) del seguente integrale doppio:

$$E[g(X, Y)] \triangleq \int_{x=0}^{\infty} \int_{y=0}^{\infty} g(x, y) f_{X_1 X_2}(x, y) dy dx < \infty$$

Riportata al caso di una coppia di variabili discrete, N e K, con lo spazio delle realizzazioni congiunte $\{(n, k) \mid n = 0, 1, 2, \dots, k = 0, 1, 2, \dots\}$ e la generica funzione $g(N, K)$, la definizione corrispondente è:

$$E[g(N, K)] \triangleq \sum_{n=0}^{\infty} \sum_{k=0}^{\infty} g(n, k) P_{N, K}(n, k) < \infty$$

La varianza della somma di due variabili aleatorie risulta:

$$\begin{aligned}
 \text{VAR}[X + Y] &\doteq E[(X + Y) - E[X + Y]]^2 \\
 &= E[((X + Y) - E[X] - E[Y])^2] \\
 &= E[X - E[X]]^2 + (Y - E[Y])^2 + 2(X - E[X])(Y - E[Y]) \\
 &= E[X - E[X]]^2 + E[Y - E[Y]]^2 + 2E[(X - E[X])(Y - E[Y])] \\
 &= \text{VAR}[X] + \text{VAR}[Y] + 2E[(X - E[X])(Y - E[Y])]
 \end{aligned}$$

Generalizzazione della formula della somma:

$$\begin{aligned}
 \text{VAR}[X_1 + X_2 + \dots + X_n] &= \\
 &= E[(X_1 + X_2 + \dots + X_n) - E[X_1 + X_2 + \dots + X_n]]^2 = \\
 &= E[(X_1 + X_2 + \dots + X_n - E[X_1] - E[X_2] - \dots - E[X_n])^2] = \\
 &= E[(X_1 - E[X_1]) + (X_2 - E[X_2]) + \dots + (X_n - E[X_n])]^2 = \\
 &= \text{VAR}[X_1] + \dots + \text{VAR}[X_n] + \\
 &\quad + 2 \sum_{i=1}^{n-1} \sum_{j=i+1}^n E[(X_i - E[X_i])(X_j - E[X_j])]
 \end{aligned}$$

Modulare la varianza attorno alla stessa media

Sia X , la v.a. che rappresenta la durata di un fenomeno e ipotizziamo di conoscere solo la durata media: $E[X]$, e di non avere informazioni sulla varianza. Di solito sceglioamo di rappresentare la durata di quel fenomeno in base ad una legge esponenziale di parametro pari al reciproco di $E[X]$.

Così facendo, implicitamente, adottiamo una varianza che è: $VAR[X] = E[X]^2$

Adesso si farà vedere che possiamo fare di meglio!

Pensiamo alla X come:

$$X \doteq X_1 + X_2 + \dots + X_n \quad \text{con} \quad X_i \approx \exp(\lambda) \quad i = 1, \dots, n \quad (\text{anche indipendenti})$$

e, sapendo che: $E[X] = n E[X_i] = n \frac{1}{\lambda}$, $VAR[X] = n VAR[X_i] = n \frac{1}{\lambda^2}$

osserviamo che, ponendo:

$$\lambda \doteq n \frac{1}{E[X]}$$

riusciamo a mantenere lo stesso valore medio $E[X] = n \frac{1}{\lambda} = n \frac{E[X]}{n}$, ma a

modificare la varianza: $VAR[X] = n \frac{E[X]^2}{n^2} = \frac{E[X]^2}{n}$ rispetto al valore che si otterrebbe

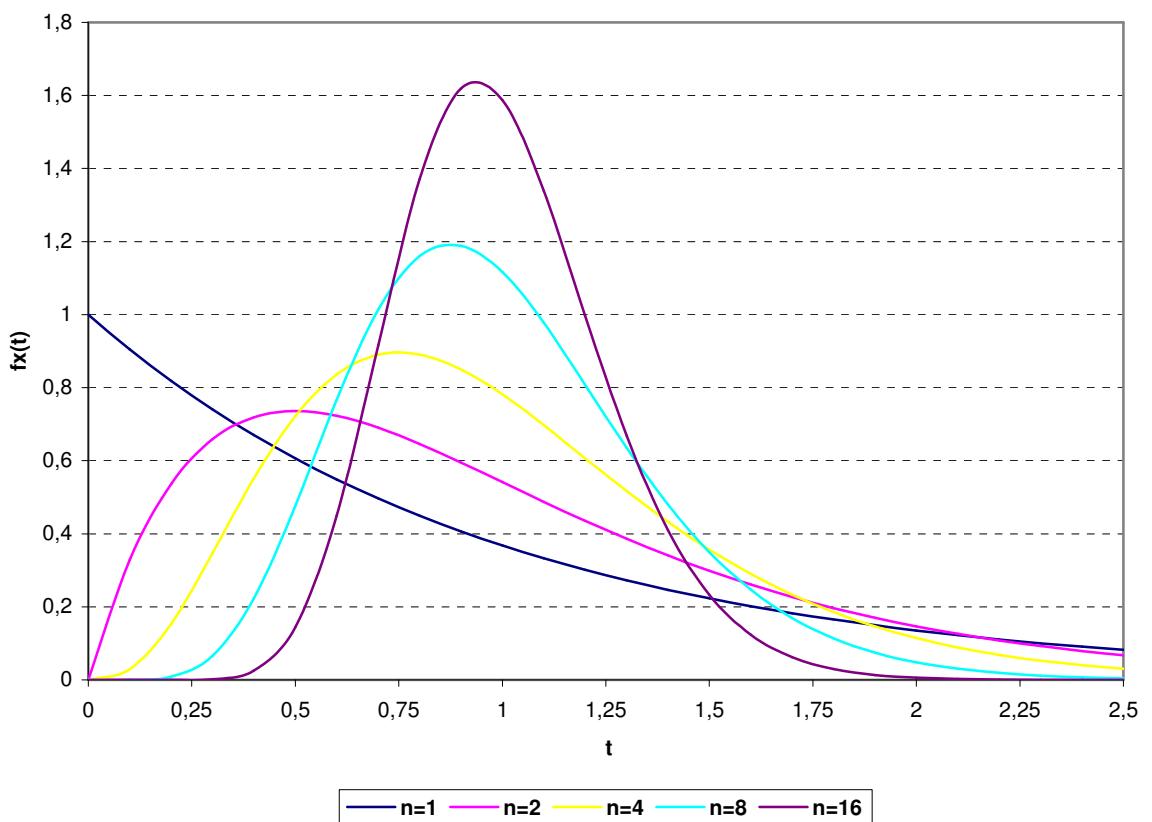
rappresentando la X con un'unica legge esponenziale.

Usando "n" come parametro di nostra scelta, possiamo far diminuire quanto vogliamo la varianza (rispetto a quella originariamente prevista dalla legge esponenziale).

Più precisamente, passiamo ad una rappresentazione del fenomeno basata su una "opportuna" distribuzione di Erlang di ordine "n", che cambia forma al crescere di "n", però mantiene sempre la stessa media.

Dunque la chiameremo Legge di Erlang MODULATA.

Densità di Erlang "modulata"



Esempio

Si supponga che un componente, una volta guasto, possa essere riparato secondo un processo di riparazione che si compone di quattro fasi in sequenza. La durata media dell'intero processo è stimata in 2 unità di tempo e non si hanno informazioni sulle durate parziali medie delle singole fasi: potrebbero corrispondere ciascuna a 0.5 unità di tempo.

Proporre una funzione densità per la variabile aleatoria "durata del processo di riparazione" e calcolare la probabilità che una riparazione duri non più di 3 unità di tempo.

Rappresentare con un grafico in Excel la suddetta densità e discutere della differenza di andamento in confronto ad una densità esponenziale caratterizzata dalla stessa media. E cosa si può dire a proposito della varianza?

La distribuzione iperesponenziale

Sia X una variabile condizionante, discreta e con un insieme finito di realizzazioni $(1, 2, \dots, n)$ e sia $(\alpha_1, \alpha_2, \dots, \alpha_n)$ un assegnamento di probabilità di tipo bernelliano generalizzato su quell'insieme. Dunque:

$$P_X(i) \hat{=} \alpha_i, \quad \alpha_i \geq 0, \quad i = 1, \dots, n \quad \text{e} \quad \sum_{i=1}^n \alpha_i = 1$$

Siano, adesso, Y_i , $i = 1, \dots, n$, "n" variabili aleatorie identicamente distribuite secondo la legge esponenziale, ciascuna però con un proprio parametro, λ_i , $i = 1, \dots, n$.

Dunque:

$$F_{Y_i}(y) \hat{=} 1 - \exp\{-\lambda_i y\}, \quad y \geq 0$$

Osservando che la densità congiunta per la coppia Y, X è:

$$\begin{aligned} f_{Y,X}(y, i) &= f_{Y|X=i}(y | i) P_X(i) \\ &= f_{Y_i}(y) P_X(i) = \alpha_i \lambda_i \exp\{-\lambda_i y\}, \end{aligned}$$

grazie alla formula della densità totale:

$$f_Y(y) = \sum_{i=1}^n f_{Y,X}(y, i)$$

si perviene ad una funzione densità (marginale) della Y che ha una forma nota come funzione densità iperesponenziale:

$$f_Y(y) = \sum_{i=1}^n \alpha_i \lambda_i \exp\{-\lambda_i y\}, \quad y \geq 0$$

e corrisponde alla combinazione convessa di densità esponenziali di parametro diverso.

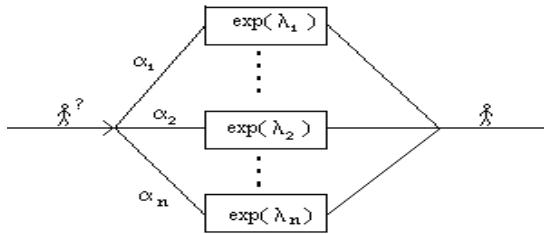
La distribuzione corrispondente risulterà, ovviamente, anch'essa data dalla combinazione convessa di funzioni di distribuzione esponenziali:

$$F_Y(y) = \int_0^y f_Y(u) du = \sum_{i=1}^n \alpha_i (1 - \exp\{-\lambda_i y\}), \quad y \geq 0$$

e può essere utilmente letta come una formula di distribuzione totale.

Visualizzazione e applicazioni

La seguente rappresentazione grafica della distribuzione iperesponenziale:



si rivela utile nelle possibili applicazioni.

Ad esempio, si consideri la situazione in cui il carico di processi (workload) che devono essere eseguiti da un sistema possa essere raggruppato in "n" classi distinte e si conosca soltanto la durata media dell'esecuzione di un processo: $1/\lambda_i$, se il processo appartiene alla classe "i". Allora, il tempo di esecuzione di un processo preso a caso fra tutti può essere modellato con la legge iperesponenziale.

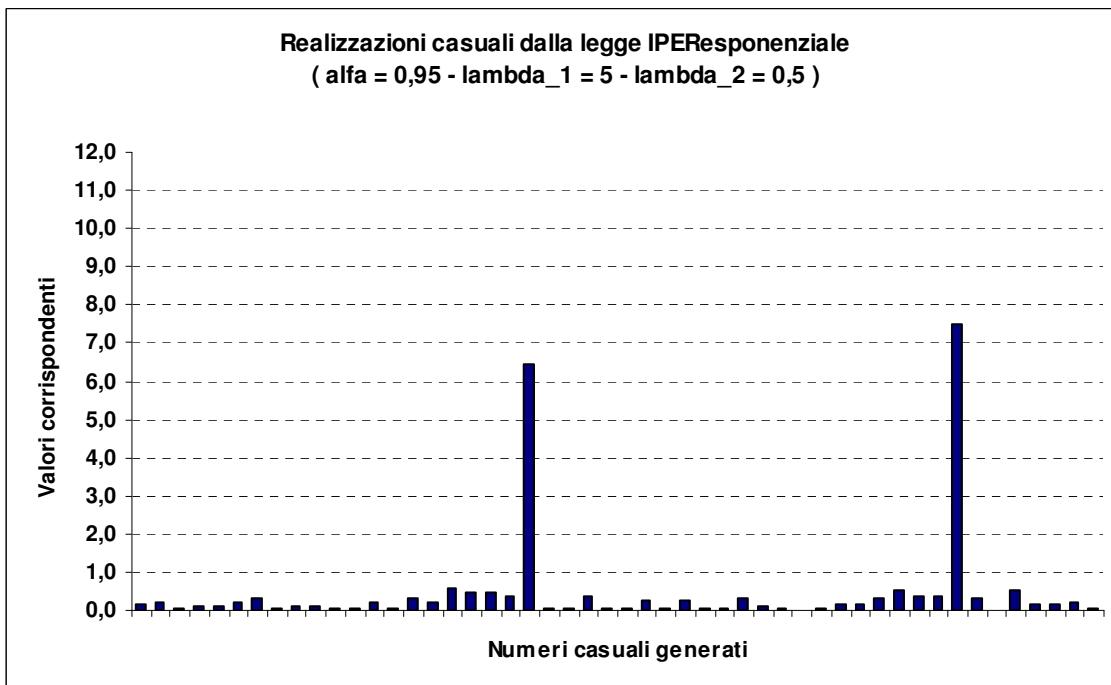
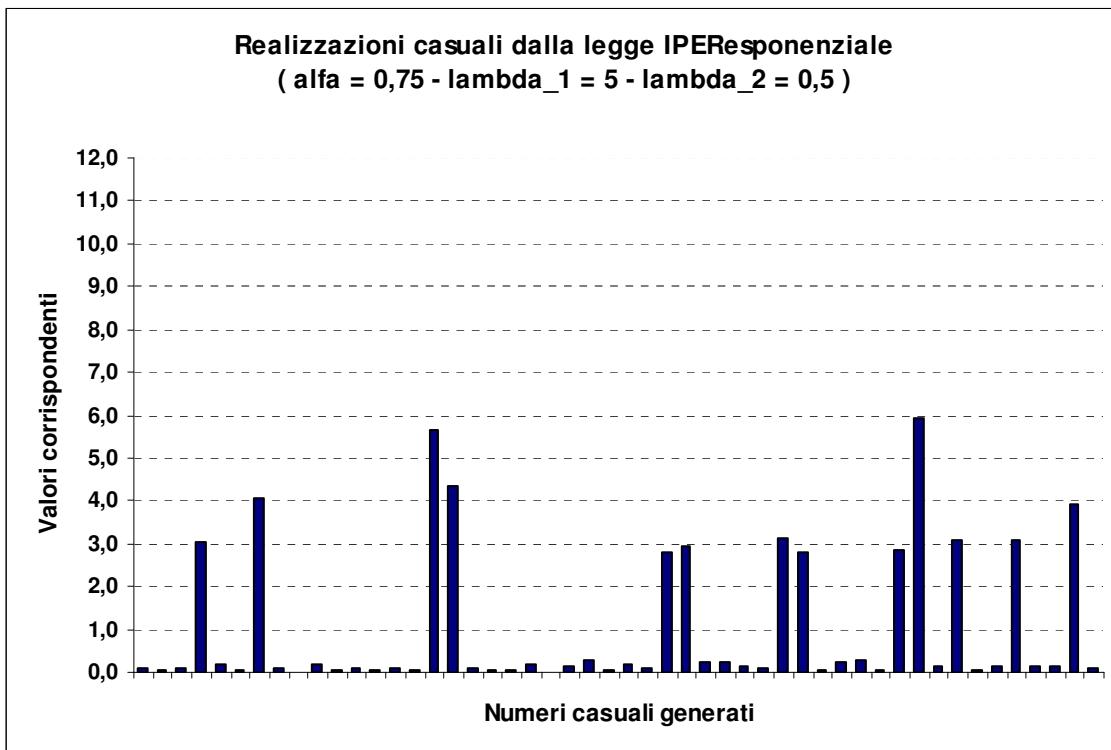
Ancora, s'immagini di limitare a due le leggi esponenziali in gioco nella figura di cui sopra e siano λ_1 e λ_2 i rispettivi parametri, con $\lambda_1 \gg \lambda_2$; in pratica, la prima esponenziale modella una durata media assai piccola se confrontata con la durata media associata alla seconda. Adesso, si aggiunga la libertà di fissare: $\alpha_1 \approx 1$ e $\alpha_2 = 1 - \alpha_1 \approx 0$. Allora, se si usa Y per rappresentare l'intervallo di tempo fra due arrivi consecutivi di "pacchetti di dati" da trasmettere su una linea, non è difficile riconoscere che lo schema in figura si presta a modellare una modalità di arrivi detta "a raffica".

In particolare, più si fissa α_1 vicino all'unità (ad es. 0.99) e più numerosa diventa la raffica, mentre, più si rende piccolo λ_2 e più tempo passa (in media) fra due raffiche consecutive.

Il nome iperesponenziale deriva dal fatto che, per essa, il rapporto: $\sqrt{VAR[Y]}/E[Y]$, detto coefficiente di variazione (Cv) risulta >1 , a fronte del fatto che il coefficiente di variazione della legge esponenziale è proprio pari a 1.

In particolare, con un $Cv >> 1$ è possibile rappresentare durate che hanno una grande varianza perché "quasi sempre molto brevi" e però "raramente assai lunghe".

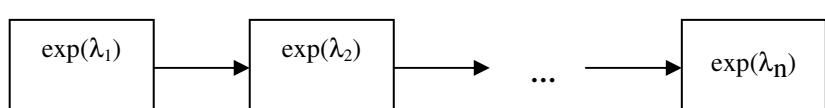
Illustrazione dell'effetto “raffica” della iperesponenziale



La distribuzione ipoesponenziale

La combinazione di fasi di durata esponenziale può essere usata anche in maniera astratta, cioè senza partire dalla descrizione della struttura di un fenomeno fisico, per generare nuove distribuzioni. Il caso più immediato è quello di un insieme di n fasi esponenziali in sequenza, ognuna di parametro $\lambda_i, i = 1, \dots, n$

e con la condizione: $\lambda_i \neq \lambda_j, i \neq j$:



Ricorrendo alla Teoria delle Trasformate di Laplace (che però è fuori programma), è possibile ricavare una nuova densità, detta ipoesponenziale ad n stadi, che generalizza la densità di Erlang di ordine n :

$$f_X(x) = \sum_{i=1}^n a_i \lambda_i \exp\{-\lambda_i x\} \quad 0 \leq x < \infty$$

e con distribuzione:

$$F_X(x) = \sum_{i=1}^n a_i (1 - \exp\{-\lambda_i x\}) \quad 0 \leq x < \infty$$

Media e varianza si calcolano immediatamente, dal diagramma, e risultano:

$$E[X] = \sum_{i=1}^n 1/\lambda_i \quad \text{e} \quad Var[X] = \sum_{i=1}^n 1/\lambda_i^2$$

Il coefficiente di variazione $\sqrt{Var[X]/E[X]}$ risulta < 1 e a ciò si deve il nome ipoesponenziale.

IL TMR con un elemento di riserva

Si vuole ricavare la distribuzione della v. a. Y , "tempo di vita del sistema", sotto le seguenti ipotesi:

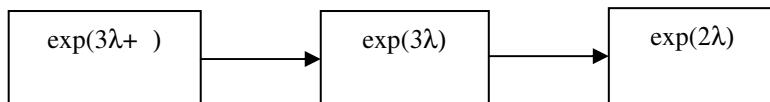
- componenti identici;
- commutazione istantanea e perfetta;
- distribuzione exp. per la v.a. $X_i \hat{=} \text{"tempo al guasto del componente } i\text{"}$:

$$F_{X_i}(t) = 1 - e^{-\lambda t}, \quad i = 1, 2, 3$$

- per il componente di riserva invece:

$$F_{X_r}(t) = \begin{cases} 1 - e^{-\lambda t} & \text{attiva} \\ 1 - e^{-\mu t} & \text{non attiva} \end{cases}, \quad \ll \lambda$$

Con il metodo delle fasi è immediato riconoscere che la funzione di distribuzione della Y è una ipoesponenziale a 3 stadi :



e quindi, ponendo per comodità: $\gamma_1 = 3\lambda+$; $\gamma_2 = 3\lambda$; $\gamma_3 = 2\lambda$

si ottiene:

$$F_Y(t) = \sum_{i=1}^3 a_i (1 - e^{-\gamma_i t}) \quad \text{dove } a_i \hat{=} \prod_{j \neq i}^3 \frac{\gamma_j}{\gamma_j - \gamma_i}$$

L'affidabilità del sistema è dunque:

$$R(t) = 1 - F_Y(t) = 1 - \sum_{i=1}^3 a_i (1 - e^{-\gamma_i t})$$

Nella pagina seguente sono raffigurati i grafici con i seguenti valori numerici:

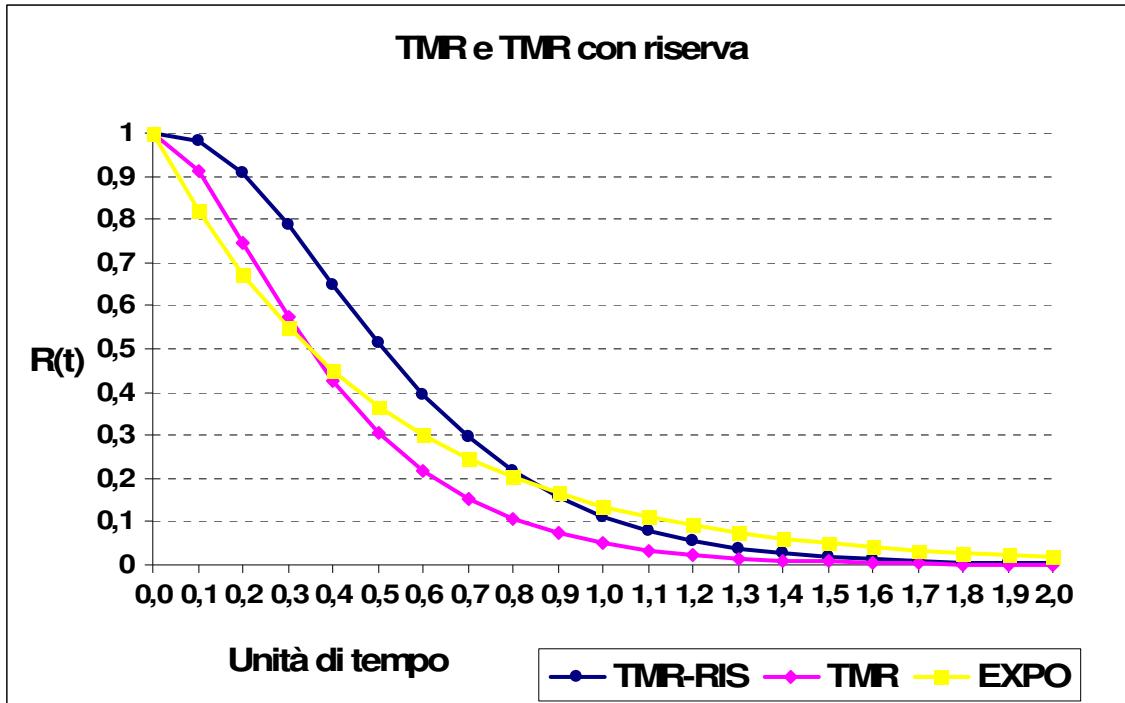
$\lambda = 2$	$\gamma_1 = 6.2$	$a_1 = 54.5$
$\mu = 0,2$	$\gamma_2 = 6$	$a_2 = -62$
	$\gamma_3 = 4$	$a_3 = 8.45$

$\lambda = 1$	$\gamma_1 = 3.2$	$a_1 = 25$
$\mu = 0,2$	$\gamma_2 = 3$	$a_2 = -32$
	$\gamma_3 = 2$	$a_3 = 8$

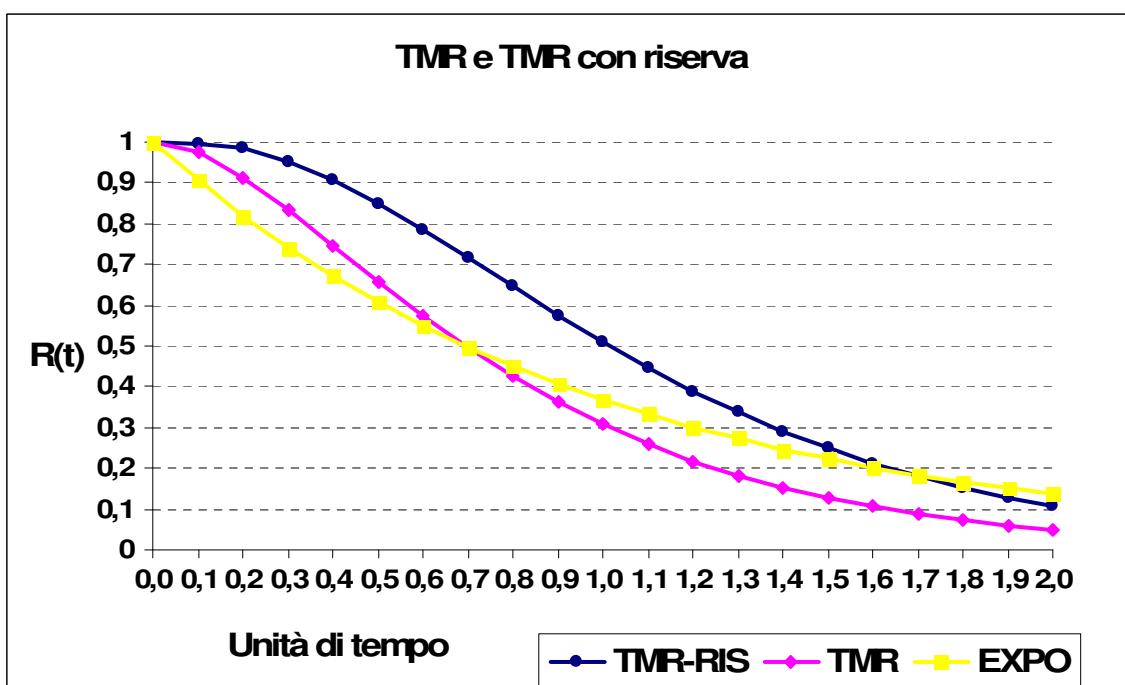
Per esercizio, si provi ad estendere il diagramma delle fasi al caso di commutazione imperfetta.

Confronto di curve di affidabilità

$$\lambda = 2, \quad = 0,2 \quad (\text{tempo}^{-1})$$



$$\lambda = 1, \quad = 0,2 \quad (\text{tempo}^{-1})$$



Esercizio di Riepilogo

Esercizio 1

Supponiamo di considerare un sistema a cui affluiscono dei messaggi. Il sistema riceve i messaggi, li inserisce in un buffer e, quando quest'ultimo è pieno, li invia su una linea di trasmissione di output. La capacità di memorizzazione del buffer è di 10 messaggi. Da un monitoraggio risulta che un singolo messaggio arriva in media ogni 2 sec.

Scegliere una funzione di distribuzioni per modellare il “tempo al riempimento del buffer” e calcolare:

1. la probabilità che il buffer si riempia in un tempo compreso fra 5 sec. e 10 sec.;
2. la probabilità che il riempimento duri più di 6 sec..

Soluzione

Siano $X_i \hat{=} \text{tempo di arrivo del messaggio } i$ con $i = 1, \dots, 10$ e siano queste variabili aleatorie tutte distribuite secondo la stessa legge esponenziale di parametro $\lambda = \frac{1}{2} \text{ sec}^{-1}$ ovvero:

$$F_{X_i}(t) = 1 - e^{-\lambda t} \text{ con } i = 1, \dots, 10.$$

Sia $X = X_1 + X_2 + \dots + X_{10} \hat{=} \text{il tempo di riempimento del buffer.}$

La funzione di distribuzione della variabile aleatoria X è la distribuzione di Erlang (di ordine 10):

$$F_X(t) = 1 - e^{-\lambda t} \sum_{i=0}^{n-1} \frac{(\lambda t)^i}{i!} \text{ con } n = 10.$$

1. La probabilità che il buffer si riempia in un tempo compreso fra 5 sec. e 10 sec.

è pari a: $P(5 \leq X \leq 10) = F_X(10) - F_X(5)$

$$F_X(10) = 1 - e^{-0,5 \cdot 10} \sum_{i=0}^9 \frac{(0,5 \cdot 10)^i}{i!}$$

$$= 1 - 0,007 \cdot (1 + 5 + 12,5 + 20,833 + 26,042 + 26,042 + 21,701 + 15,502 + 9,688 + 5,382)$$

$$\approx 0,03183$$

$$F_X(5) = 1 - e^{-0,5 \cdot 5} \sum_{i=0}^9 \frac{(0,5 \cdot 5)^i}{i!} =$$

$$= 1 - 0,82 \cdot (1 + 2,5 + 3,125 + 2,604 + 1,628 + 0,814 + 0,339 + 0,121 + 0,038 + 0,011)$$

$$\cong 0,00028$$

$$F_X(10) - F_X(5) = 0,03183 - 0,00028 = 0,03155$$

2. La probabilità che il riempimento duri più di 6 sec. è pari a:

$$P(X > 6) = 1 - P(X \leq 6) = 1 - F_X(6)$$

$$F_X(6) = 1 - e^{-0,5 \cdot 6} \sum_{i=0}^9 \frac{(0,5 \cdot 6)^i}{i!}$$

$$= 1 - 0,0498 \cdot (1 + 3 + 4,5 + 4,5 + 3,375 + 2,025 + 1,013 + 0,434 + 0,163 + 0,054)$$

$$\cong 0,0011$$

$$1 - F_X(6) = 0,9989$$

Esercizio 2

Un sistema prevede che una CPU esegua processi allocati su 2 *ready queues* distinte ed indipendenti, svuotando prima l'una e poi l'altra coda (ciclicamente). Da un monitoraggio HW risulta che la durata media di “attività continua” della CPU sulla prima coda è pari a 40msec., mentre sulla seconda è 80msec.. Tralasciando di considerare gli intervalli di ozio forzato della CPU, si chiede di:

1. ricavare la funzione di distribuzione della durata del ciclo di attività della CPU;
2. calcolare la probabilità che la lunghezza di un ciclo non superi i 50 msec..

Soluzione

1. Il ciclo di attività della CPU è praticamente il tempo necessario per svuotare le due code. Sia $X_i \hat{=} \text{tempo necessario per svuotare la coda } i$ con $i = 1,2$ dove entrambe le variabili aleatorie sono distribuite secondo una legge esponenziale

$F_{X_i} = 1 - e^{-\lambda_i t}$, ma di parametro $\lambda_1 = 1/40 \text{ msec} = 25 \text{ Hz}$ e $\lambda_2 = 1/80 \text{ msec} = 12,5 \text{ Hz}$, rispettivamente.

Allora $X = X_1 + X_2 \hat{=} \text{ la durata aleatoria del ciclo di attività della CPU è una variabile aleatoria la cui distribuzione si ricava come segue:}$

$$F_X(t) = P(X_1 + X_2 \leq t) = \int_{x_2=0}^t F_{X_1}(t-x_2) f_{X_2}(x_2) dx_2 .$$

Dunque

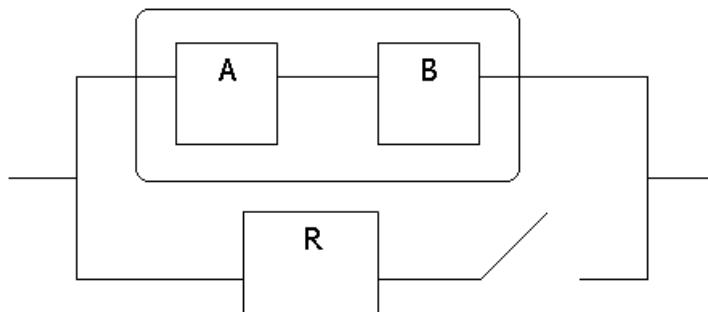
$$\begin{aligned} F_X(t) &= \int_{x_2=0}^t [1 - e^{-\lambda_1(t-x_2)}] \lambda_2 e^{-\lambda_2 x_2} dx_2 = \lambda_2 \int_{x_2=0}^t e^{-\lambda_2 x_2} dx_2 - \lambda_2 e^{-\lambda_1 t} \int_{x_2=0}^t e^{-(\lambda_2-\lambda_1)x_2} dx_2 \\ &= -\frac{\lambda_2}{\lambda_2} e^{-\lambda_2 x_2} \Big|_0^t + \frac{\lambda_2}{\lambda_2 - \lambda_1} e^{-(\lambda_2 - \lambda_1)x_2} \Big|_0^t \\ &= -\left(e^{-\lambda_2 t} - e^{-\lambda_2 0}\right) + \frac{\lambda_2}{\lambda_2 - \lambda_1} e^{-\lambda_1 t} \left(e^{-(\lambda_2 - \lambda_1)t} - e^{-(\lambda_2 - \lambda_1)0}\right) \\ &= 1 - e^{-\lambda_2 t} + \frac{\lambda_2}{\lambda_2 - \lambda_1} e^{-\lambda_1 t} \left(e^{-(\lambda_2 - \lambda_1)t} - 1\right) = 1 - e^{-\lambda_2 t} + \frac{\lambda_2}{\lambda_2 - \lambda_1} e^{-\lambda_2 t} - \frac{\lambda_2}{\lambda_2 - \lambda_1} e^{-\lambda_1 t} \\ &= 1 - e^{-\lambda_2 t} \left(1 - \frac{\lambda_2}{\lambda_2 - \lambda_1}\right) - \frac{\lambda_2}{\lambda_2 - \lambda_1} e^{-\lambda_1 t} = 1 - e^{-\lambda_2 t} \left(\frac{\lambda_2 - \lambda_1 - \lambda_2}{\lambda_2 - \lambda_1}\right) - \frac{\lambda_2}{\lambda_2 - \lambda_1} e^{-\lambda_1 t} \\ &= 1 + \frac{\lambda_1}{\lambda_2 - \lambda_1} e^{-\lambda_2 t} - \frac{\lambda_2}{\lambda_2 - \lambda_1} e^{-\lambda_1 t} = 1 - \left(\frac{\lambda_2}{\lambda_2 - \lambda_1} e^{-\lambda_1 t} - \frac{\lambda_1}{\lambda_2 - \lambda_1} e^{-\lambda_2 t}\right). \end{aligned}$$

2. la probabilità che la lunghezza di un ciclo non superi i 50msec. è data da:

$$\begin{aligned} P(X \leq 50 \text{ msec}) &= F_X(50 \text{ msec}) = 1 - \left(\frac{\lambda_2}{\lambda_2 - \lambda_1} e^{-\lambda_1 t} - \frac{\lambda_1}{\lambda_2 - \lambda_1} e^{-\lambda_2 t}\right) \\ &= 1 - \left(-1 \cdot e^{-12,5 \cdot 0,05} + 2e^{-25 \cdot 0,05}\right) = 1 - (-0,2865 + 1,0705) = 1 - 0,784 = 0,216 \end{aligned}$$

Esercizio 3

Due apparecchi elettronici indipendenti, identici e non riparabili, sono organizzati in serie. Ad essi è affiancato un terzo apparecchio che funge da riserva pronta, come mostrato in Figura.



Sono noti i tempi medi al guasto di tutti i componenti ovvero 10.000 ore per A , 10.000 ore per B e 8.000 ore per il componente di riserva R . Si chiede di:

1. giustificare un modello di affidabilità comune per tutti i componenti e con quello calcolare il tempo medio al guasto del sottosistema seriale;
2. assumendo che la commutazione riesca sempre e precisando ulteriori ipotesi, ricavare media e varianza del tempo di vita del sistema;
3. impostare la formula che definisce la distribuzione del “tempo di vita del sistema”.

Soluzione

1. Si considerino le seguenti variabili aleatorie:

$X_A \hat{=} \text{tempo al guasto del componente } A;$

$X_B \hat{=} \text{tempo al guasto del componente } B;$

$X_R \hat{=} \text{tempo al guasto del componente di riserva } R$

ciascuna distribuita con legge esponenziale di parametro

$$\lambda_A = \lambda_B = \frac{1}{10000} = 10^{-4} \text{ e } \lambda_R = \frac{1}{8000} = 1,25 * 10^{-4}, \text{ rispettivamente.}$$

Il tempo medio al guasto del sottosistema seriale è calcolabile a partire dalla formula dell'affidabilità di detto sistema ovvero:

$$R_{AB}(t) = R_A(t) \cdot R_B(t) = [1 - F_{X_A}(t)] \cdot [1 - F_{X_B}(t)] = [1 - 1 + e^{-\lambda_A t}] \cdot [1 - 1 + e^{-\lambda_B t}] = e^{-(\lambda_A + \lambda_B)t}$$

da cui

$$\begin{aligned} E[X_{AB}] &= \int_{u=0}^{\infty} R_{AB}(u) du = \int_{u=0}^{\infty} e^{-(\lambda_A + \lambda_B)u} du \\ &= -\frac{1}{\lambda_A + \lambda_B} e^{-(\lambda_A + \lambda_B)u} \Big|_0^{\infty} = -\frac{1}{2 \cdot 10^{-4}} e^{-2 \cdot 10^{-4} \cdot \infty} + \frac{1}{2 \cdot 10^{-4}} e^{-2 \cdot 10^{-4} \cdot 0} = 5000 \end{aligned}$$

2. Assumendo che i componenti siano indipendenti, che la commutazione riesca sicuramente e che questa si realizza in tempo trascurabile, per ricavare la media e varianza del tempo di vita del sistema si definisca la seguente variabile aleatoria:

$Y \hat{=} \text{tempo di vita del sistema.}$

Allora

$$E[Y] = E[X_{AB} + X_R] = E[X_{AB}] + E[X_R]$$

$$\frac{1}{\lambda_{AB}} + \frac{1}{\lambda_R} = \frac{1}{\lambda_A + \lambda_B} + \frac{1}{\lambda_R} = 5000 + 8000 = 13000$$

mentre

$$VAR[Y] = VAR[X_{AB} + X_R] = VAR[X_{AB}] + VAR[X_R] + \underbrace{2E[(X_{AB} - E(X_{AB}))(X_R - E(X_R))]}_{=0}$$

3. Per impostare la formula che definisce la distribuzione del “tempo di vita del sistema”, si assume che la commutazione riesca sicuramente e si realizzi in tempo trascurabile. Sia

$Y \hat{=} \text{tempo al guasto del sistema} = X_{AB} + X_R.$

Le funzioni di distribuzione e densità sono rispettivamente

$$\begin{cases} F_{X_{AB}}(t) = 1 - e^{-(\lambda_A + \lambda_B)t} \\ f_{X_{AB}}(t) = (\lambda_A + \lambda_B)e^{-(\lambda_A + \lambda_B)t} \end{cases}$$

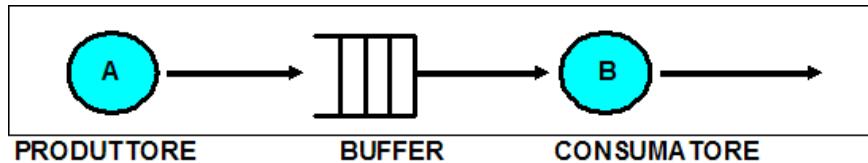
$$\begin{cases} F_{X_R}(t) = 1 - e^{-\lambda_R t} \\ f_{X_R}(t) = \lambda_R e^{-\lambda_R t} \end{cases}$$

Dai risultati ottenuti sulla distribuzione della somma di variabili aleatorie indipendenti, si ha:

$$F_Y(t) = \int_{x_1=0}^t F_{X_R}(t-x_1) f_{X_{AB}}(x_1) dx_1 = \int_{x_1=0}^t (1 - e^{-\lambda_R(t-x_1)}) [(\lambda_A + \lambda_B) e^{-(\lambda_A + \lambda_B)x_1}] dx_1$$

Esercizio 4

In un sistema informatico, due risorse particolari (A e B) sono in relazione secondo il ben noto protocollo produttore-consumatore (o cliente-servente). Poiché lavorano in maniera asincrona, è previsto un buffer intermedio per accumulare temporaneamente gli oggetti prodotti da A e destinati ad essere consumati da B (come in figura). Il buffer ha una disponibilità limitata di posti, pari a “ k ”. Quando il buffer è pieno il produttore si blocca e riprenderà la produzione di oggetti solo dopo che sarà intervenuto il consumatore a svuotare il buffer stesso.



La risorsa consumatore (B) lavora alla seguente maniera: in corrispondenza di istanti di tempo scelti in maniera completamente casuale, provvede a consumare uno dopo l’altro tutti gli oggetti che trova nel buffer, fino allo svuotamento completo. Il valore atteso del tempo impiegato a consumare il singolo oggetto è pari a $1/\lambda_B$ e non ci sono tempi morti tra un consumo e il successivo. Prima di cominciare lo svuotamento, se non trova il buffer pieno e il produttore bloccato, il consumatore si preoccupa di inibire comunque la fase di produzione mandando un messaggio al produttore.

Nelle ipotesi appena descritte, il problema dello svuotamento del buffer consiste nel ricavare la funzione di distribuzione della variabile aleatoria “tempo necessario a

realizzare un’operazione di svuotamento” del buffer. Oppure, in maniera più limitata e con un approccio solo numerico, si può ricavare la probabilità che il tempo di svuotamento sia superiore ad un valore prefissato.

La risoluzione del problema richiede una fase preliminare dedicata ad individuare un possibile modello di occupazione del buffer. Nella fase successiva, adottando una legge esponenziale per la variabile aleatoria che rappresenta il “tempo di consumo del singolo oggetto” si farà ricorso alla legge di Erlang di ordine “n” per modellare la durata dello svuotamento, sotto la condizione che siano proprio n gli oggetti trovati nel buffer dal consumatore. Da qui, applicando la formula della distribuzione totale con variabile condizionante corrispondente alla variabile aleatoria “numero di oggetti trovati nel buffer dal consumatore” verrà ricavato il risultato cercato.

FASE PRELIMINARE: Il modello di occupazione del buffer

Per giustificare un possibile modello di occupazione del buffer, ovvero la funzione di ripartizione della variabile aleatoria (N) che rappresenta il numero di oggetti trovati dal consumatore (che agisce come un osservatore casuale!) si ragionerà alla seguente maniera. Il consumatore compie l’esperimento aleatorio che consiste nel ripetere l’ispezione del singolo posto nel buffer. Il risultato della singola ispezione visiva è binario: posto occupato da un oggetto da consumare oppure posto libero. A questo punto, se escludiamo il fenomeno di “frammentazione” dell’occupazione, assumendo che non ci possono essere posti liberi fra due posti occupati, possiamo dire che il consumatore trova un posto occupato con probabilità ρ , $0 < \rho < 1$, e un posto libero con probabilità $1 - \rho$. Allora, la probabilità che il buffer contenga n oggetti equivale alla probabilità che il primo successo (posto libero) si ottenga al tentativo (ispezione visiva) $n+1$.

Dunque, la legge geometrica estesa anche al risultato nullo (buffer vuoto!) scaturisce in maniera immediata come punto di partenza per raffinare il modello di occupazione del buffer.

A tal fine essa viene qui ripresa:

$$P_n = (1 - \rho)\rho^n, \quad n = 0, 1, \dots, k, k + 1, \dots$$

osservando che, però, nella formulazione originaria, essa prevede un numero di ispezioni infinitamente numerabile. In altri termini, un buffer ideale con un numero di posti illimitato!

L'adattamento al caso di nostro interesse si realizza sfruttando il concetto di probabilità condizionata, procedendo alla seguente maniera.

Evento di interesse: *esattamente n utenti presenti | al più k possono essere accettati*
 $\hat{=} n | n \leq k$

$$P_{n \leq k} = \sum_{n=0}^k (1-\rho)\rho^n$$

Probabilità evento condizionante:

$$P_{n|n \leq k} = \frac{P_n}{P_{n \leq k}} = \frac{(1-\rho)\rho^n}{\sum_{n=0}^k (1-\rho)\rho^n}, \quad n = 0, 1, \dots, k$$

Derivazione della formula finale per $P_{n|n \leq k}$, $n = 0, 1, \dots, k$

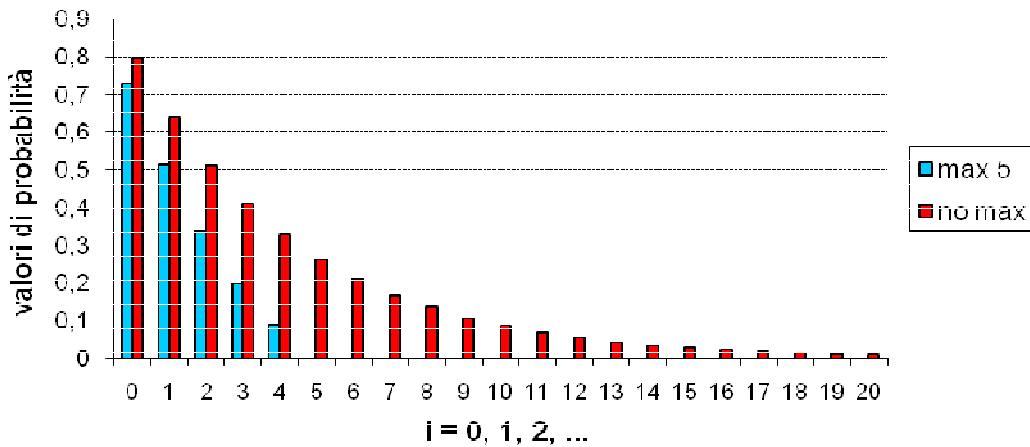
Si sviluppa:

$$\begin{aligned} \sum_{n=0}^k \rho^n &= \sum_{i=0}^{\infty} \rho^n - \sum_{n=k+1}^{\infty} \rho^n \quad \text{con } \rho < 1 \text{ e } j \hat{=} n - (k+1) \Rightarrow \\ &= \frac{1}{1-\rho} - \sum_{j=0}^{\infty} \rho^{k+1} \rho^j = \frac{1}{1-\rho} - \frac{\rho^{k+1}}{1-\rho} = \frac{1-\rho^{k+1}}{1-\rho} \end{aligned}$$

E si ottiene:

$$P_{n|n \leq k} = \frac{\rho^n}{\sum_{n=0}^k \rho^n} = \frac{(1-\rho)}{1-\rho^{k+1}} \rho^n, \quad n = 0, 1, \dots, k \quad (\textbf{Geometrica Troncata})$$

Probabilità (oggetti nel buffer > i), con $p = 0.8$



PRIMA FASE: la distribuzione condizionata del tempo allo svuotamento

Adesso, adottando una legge esponenziale di parametro λ per la variabile aleatoria (S) che rappresenta il “tempo di consumo del singolo oggetto” si vedrà come si può ricorrere alla legge di Erlang di ordine “ n ” per modellare la durata dello svuotamento, sotto la condizione che siano proprio “ n ” gli oggetti trovati nel buffer dal consumatore. Anzitutto si assume che, se il consumatore trova “ n ” oggetti nel buffer quando si accinge ad avviare le operazioni di consumo, “ n ” è inteso come realizzazione indipendente della variabile aleatoria N , distribuita secondo una legge una geometrica troncata. Allora vorrà dire che, per svuotare completamente il buffer, il consumatore impiegherà un tempo pari proprio alla somma di “ n ” realizzazioni esponenziali della variabile aleatoria S . Questo perché non ammettiamo tempi morti tra un consumo e il successivo. Se aggiungiamo, a questo punto, l’ipotesi che i singoli consumi siano operazioni identiche ma di durata indipendente una dall’altra e tutte insieme, allora risultano verificate tutte le ipotesi necessarie e sufficienti a concludere che la variabile aleatoria (condizionata) “ $W|N=n$ ” – intesa come tempo allo svuotamento in presenza di n oggetti - segue una legge di Erlang di ordine “ n ”:

$$F_{W|N}(t|n) = 1 - e^{-\lambda t} \sum_{i=0}^{n-1} \frac{(t)^i}{i!}; \text{ dove, } F_{W|N=0}(t) = 1 \text{ e } \lambda \triangleq 1/E[S] \text{ (ritmo di consumo)}$$

$$\text{Con } P(N = n) = \frac{(1-\rho)}{1-\rho^{k+1}} \rho^n, \quad n = 0, 1, \dots, k \Rightarrow F_{W|N=0}(t) \cdot P(N = 0) = 1 \cdot (1-\rho)$$

SECONDA FASE: applicazione della formula della distribuzione totale

Per fissare le idee anche su un esempio numerico, s'immagini di voler calcolare la probabilità che il tempo (incondizionato) allo svuotamento del buffer sia superiore a 2 u.t., assumendo un tempo medio del singolo consumo pari a 1 ($E[S]=1$).

A titolo di esempio, si adotterà una geometrica troncata al valore $k = 4$ e con parametro “ ρ ” fissato al valore 0.75. Dunque:

$$P(N=0)=0,365; P(N=1)=0,274 P(N=2)=0,206 ; P(N=3)=0,155$$

Applicando la seguente formula della distribuzione totale:

$$F_W^{(m)}(t) \hat{=} \Pr[W \leq t] = \sum_{n=0}^4 \Pr[W \leq t | N = n] \cdot \Pr[N = n] = \sum_{n=0}^4 F_{W|N}(t | n) \cdot \Pr[N = n]$$

Si ottiene:

$$\begin{aligned} F_W(t) &= \left[1 - e^{-\mu t} \right] \cdot P(N=1) + \left[1 - e^{-\mu t} \cdot \left(1 + \frac{(\mu t)^1}{1!} \right) \right] \cdot P(N=2) + \\ &+ \left[1 - e^{-\mu t} \cdot \left(1 + \frac{(\mu t)^1}{1!} + \frac{(\mu t)^2}{2!} \right) \right] \cdot P(N=3) \end{aligned}$$

Da cui la probabilità che il tempo di svuotamento superi le 2 u.t.:

$$\begin{aligned} 1 - F_X^{(m)}(t = 2) &= 1 - [0.865 \cdot 0.274 + (1 - 0.135 \cdot 3) \cdot 0.206 + \\ &+ (1 - 0.135 \cdot 5) \cdot 0.155] = 0.59 \end{aligned}$$

Valore atteso totale e curva di regressione

In analogia con la formula della distribuzione totale, verrà adesso ricavata una formula "di tipo totale" per il valore atteso di una variabile aleatoria (continua e non negativa).

Partendo dalla definizione di valore atteso:

$$E[Y] \hat{=} \int_{y=0}^{\infty} y f_Y^{(m)}(y) dy$$

si esprime la densità (marginale) f mediante la formula della densità totale:

$$= \int_{y=0}^{\infty} y \int_{x=0}^{\infty} f_{Y|X}(y/x) f_X^{(m)}(x) dx dy = \int_{x=0}^{\infty} \int_{y=0}^{\infty} y f_{Y|X}(y/x) dy f_X^{(m)}(x) dx$$

e riconoscendo che è naturale definire

come valore atteso condizionato il seguente: $E[Y|X] \hat{=} \int_{y=0}^{\infty} y f_{Y|X}(y/x) dy$

si ottiene la formula del valore atteso totale:

$$E[Y] = \int_{x=0}^{\infty} E[Y|X] f_X^{(m)}(x) dx$$

Con una variabile condizionante, X, discreta:

$$E[Y] = \sum_{x=0}^{\infty} E[Y|X] P_X(x)$$

OSSERVAZIONE:

La formula del valore atteso condizionato non definisce un numero, ma una funzione sullo spazio delle realizzazioni (continue e non negative) della X:

$$E[Y|X=x], \quad 0 \leq x < \infty.$$

Tale funzione, che descrive l'andamento del valore atteso della Y al variare della x, è detta curva di regressione e si può ricavare, in linea di principio, a partire dalla conoscenza della densità condizionata, $f_{Y|X}(y|x)$ (o della congiunta, $f_{Y,X}(y,x)$).

Nelle applicazioni pratiche è compito della Statistica stimare i parametri della curva di regressione, a partire da osservazioni sperimentali della coppia (X,Y).

Concetto e formula della covarianza

Si definisce covarianza di due variabili aleatorie, X e Y, e si indica $COV(X,Y)$ il seguente valore atteso:

$$E[(X - E[X])(Y - E[Y])] \doteq COV(X,Y)$$

utile perché offre un'indicazione (sia pure abbastanza limitata) di come X e Y varino l'una relativamente all'altra. Infatti, se piccole realizzazioni della X tendono ad essere associate a piccole realizzazioni della Y e, viceversa, grandi realizzazioni dell'una a grandi dell'altra, allora le grandezze (aleatorie) $(X - E[X])$ e $(Y - E[Y])$ avranno lo stesso segno (positivo) e risulterà: $COV(X,Y) > 0$. Ragionando analogamente, ci si aspetterà covarianza negativa nel caso in cui a piccole realizzazioni dell'una variabile aleatoria corrispondano grandi realizzazioni dell'altra.

La formula di calcolo della covarianza è:

$$\begin{aligned} COV(X,Y) &\doteq E[(X - E[X])(Y - E[Y])] \\ &= E[XY - YE[X] - XE[Y] + E[X]E[Y]] \\ &= E[XY] - E[X]E[Y] \end{aligned}$$

Correlazione di una coppia di variabili

Le variabili aleatorie X e Y sono dette correlate quando risulta $COV(X,Y) \neq 0$ e, viceversa, quando risulta $COV(X,Y) = 0$ X e Y sono dette non correlate.

In particolare, se sono indipendenti vale il risultato: $E[XY] = E[X]E[Y]$ e quindi le due variabili sono pure non correlate. Ma non vale il viceversa, nel senso che le due variabili possono essere a covarianza nulla senza essere pure indipendenti!

In entrambi i casi, di indipendenza o solo di correlazione nulla, si noti che risulta valida la seguente implicazione sulla varianza della somma di X e Y:

$$\begin{aligned} VAR[X + Y] &= VAR[X] + VAR[Y] + 2COV(X,Y) && (*) \\ &= VAR[X] + VAR[Y] && (\text{additività della varianza}) \end{aligned}$$

Per completezza, si tenga anche presente che la proprietà di additività si generalizza con la seguente:

$$VAR\left[\sum_{i=1}^n c_i X_i\right] = \sum_{i=1}^n c_i^2 VAR[X_i]$$

per una successione di variabili X_i e una successione di costanti arbitrarie c_i .

Il coefficiente di Pearson

Poiché dalla (*) si evince che la covarianza ha le dimensioni di una varianza, non è difficile accettare la seguente misura di correlazione:

$$\rho(X, Y) \triangleq \frac{COV(X, Y)}{\sqrt{VAR[X] VAR[Y]}}$$

adimensionale e indipendente dalla scala delle quantità in gioco, nota come coefficiente di Pearson. Adesso si farà vedere che risulta: $-1 \leq \rho(X, Y) \leq 1$ e che tale coefficiente è una misura del grado di linearità della correlazione fra le variabili X e Y.

Proprietà del coefficiente di Pearson (CdP).

Proprietà 1

Si dimostra che: $-1 \leq \rho(X, Y) \leq 1$

Prova:

Siano W una varia aleatoria $E[W^2] \neq 0$ e Z una variabile aleatoria $E[Z^2] \neq 0$

allora

$$(aW - Z)^2 \geq 0 \quad (a \in \mathbb{R}) \Rightarrow E[(aW - Z)] \geq 0.$$

Ma

$$E[(aW - Z)^2] = a^2 E[W^2] - 2a \cdot E[W \cdot Z] + E[Z^2] \geq 0.$$

Definendo

$$a \triangleq E[W \cdot Z] / E[W^2]$$

per sostituzione si ottiene

$$\begin{aligned} & \frac{(E[W \cdot Z])^2}{(E[W^2])^2} E[W^2] - 2a \cdot \frac{(E[W \cdot Z])^2}{E[W^2]} + E[Z^2] \geq 0 \\ & \Rightarrow -\frac{(E[W \cdot Z])^2}{E[W^2]} + E[Z^2] \geq 0 \quad \Rightarrow \quad \frac{(E[W \cdot Z])^2}{E[W^2] \cdot E[Z^2]} \leq 1. \end{aligned}$$

Con

$$W \doteq X - \bar{X} \text{ e } Z \doteq Y - \bar{Y} \Rightarrow -1 \leq \rho(X, Y) \leq +1.$$

Proprietà 2

Si dimostra che: $|\rho(X, Y)| = 1 \Leftrightarrow Y = \alpha + \beta X$, con α e β reali e $\beta \neq 0$

Prova: (si dimostra solo l'implicazione \Leftarrow)

Dalla relazione: $Y = \alpha + \beta X$

$$\begin{aligned} \text{si ha:} \quad COV[X, Y] &= E[XY] - E[X]E[Y] \\ &= \alpha E[X] + \beta E[X^2] - E[X](\alpha + \beta E[X]) \\ &= \beta VAR[X] \end{aligned}$$

e dalla: $X = \beta^{-1}Y - \alpha$

allo stesso modo: $COV[X, Y] = \beta^{-1}VAR[Y]$

Dunque:

$$COV[X, Y]^2 = VAR[X]VAR[Y]$$

A partire dalla precedente si può, allora, concludere che:

$$\rho(X, Y) = \begin{cases} -1 & Y = -\beta X + \alpha, \quad \beta > 0 \\ 0 & X, Y \text{ non correlate} \\ +1 & Y = \beta X + \alpha, \quad \beta > 0 \end{cases}$$

La retta di regressione

La retta di regressione è quella particolare curva di regressione che si ottiene ponendo:

$$E[Y | X = x] \hat{=} a \cdot x + b, \quad 0 \leq x < \infty$$

ed è comunemente usata nell'analisi (statistica) della dipendenza della variabile aleatoria Y dalla X, dopo aver stimato i coefficienti reali a e b a partire dalle realizzazioni sperimentali della coppia X,Y.

TEOREMA (*sui coefficienti della retta di regressione*)

Quale che sia la forma delle funzioni di distribuzione della X e della Y, i coefficienti a e b possono essere ricavati dalle seguenti formule:

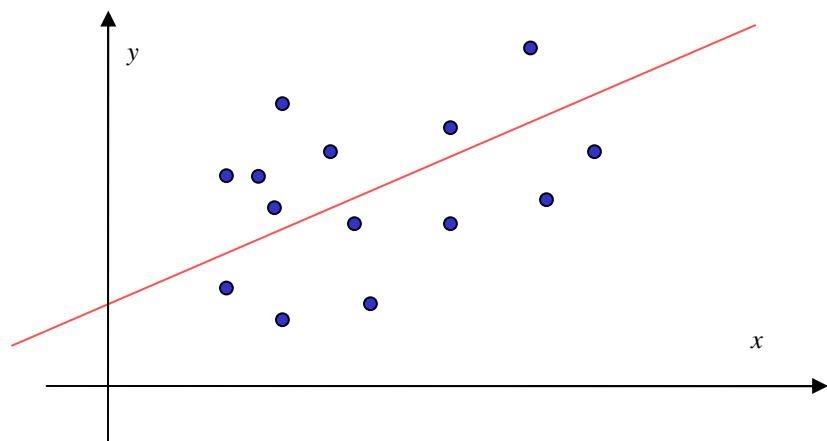
$$a = \rho \frac{\sqrt{VAR[Y]}}{\sqrt{VAR[X]}} = \frac{COV(X, Y)}{VAR[X]}, \quad b = E[Y] - \frac{COV(X, Y)}{VAR[X]} E[X];$$

dove ρ è il coefficiente di correlazione di Pearson.

Dimostrazione (in seguito)

Stima dei coefficienti

Qui si farà vedere come possano essere determinati i parametri (a e b) della retta di regressione, a partire da un insieme di coppie $\{(x_1, y_1), (x_2, y_2), \dots, (x_n, y_n)\}$ di realizzazioni congiunte delle variabili aleatorie X e Y . L'idea è che la retta di regressione debba essere proprio quella che, nel piano euclideo, passa "il più possibile vicino" ai punti corrispondenti alle coppie di realizzazioni $\{(x_1, y_1), (x_2, y_2), \dots, (x_n, y_n)\}$. La figura seguente illustra l'idea:



Per formalizzare il concetto di “il più possibile vicino”, si conviene di cercare i parametri a e b tali che risulti minima la somma dei quadrati degli scostamenti dei valori y_1, y_2, \dots, y_n osservati rispetto ai

$$\text{valori sulla retta stessa: } S = \sum_{i=1}^n [y_i - (ax_i + b)]^2.$$

Considerando, dunque, la precedente somma (S) come funzione dei due parametri a e b , si deriva prima rispetto all’uno e poi rispetto all’altro:

$$\begin{cases} \frac{\partial S(a,b)}{\partial b} = -2 \cdot \sum_{i=1}^n (y_i - b - ax_i) = 0 \\ \frac{\partial S(a,b)}{\partial a} = -2 \cdot \sum_{i=1}^n (y_i - b - ax_i) \cdot x_i = 0 \end{cases}; \quad \text{da cui:}$$

$$\begin{cases} n \cdot b + a \cdot \sum_{i=1}^n x_i - \sum_{i=1}^n y_i = 0 \\ b \cdot \sum_{i=1}^n x_i + a \cdot \sum_{i=1}^n x_i^2 - \sum_{i=1}^n x_i y_i = 0 \end{cases}$$

Risolvendo rispetto ad a e b si ricavano le stime ai minimi quadrati (denotate come “ \hat{a} ” e “ \hat{b} ”):

$$\hat{a} = \frac{n \cdot \sum_{i=1}^n x_i \cdot y_i - \sum_{i=1}^n x_i \sum_{i=1}^n y_i}{n \cdot \sum_{i=1}^n x_i^2 - \left(\sum_{i=1}^n x_i \right)^2} \quad \text{e} \quad \hat{b} = \frac{\sum_{i=1}^n y_i}{n} - \hat{a} \cdot \frac{\sum_{i=1}^n x_i}{n}, \quad (\text{r3})$$

che risultano in accordo con le formule già ricavate: $a = \frac{COV(X,Y)}{VAR[X]}$ e $b = E[Y] - a \cdot E[X]$.

Esempi di rette di regressione

Due esempi numerici di costruzione di rette di regressione sono illustrati nelle prossime due figure e sono tratti da un caso reale di analisi della correlazione presso un terminale marittimo per container, dove erano state individuate due sottoaree separate di stoccaggio dei container sul piazzale che potevano ospitare gruppi di container che arrivavano con la stessa nave e con una seconda nave, ancora comune, erano destinati a ripartire. In tal caso, sia i tempi di giacenza nelle due sottoaree sia i rispettivi livelli di occupazione dovevano risultare dipendenti, come confermato dalle rette di regressione ricavate con i dati della pagina seguente.

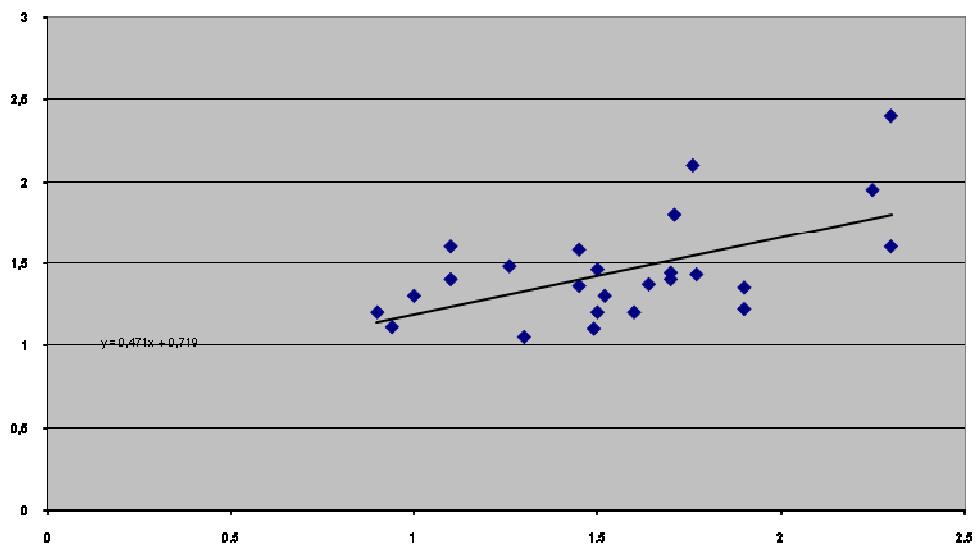


Fig. 1. Retta di regressione per i tempi di giacenza (settimane)

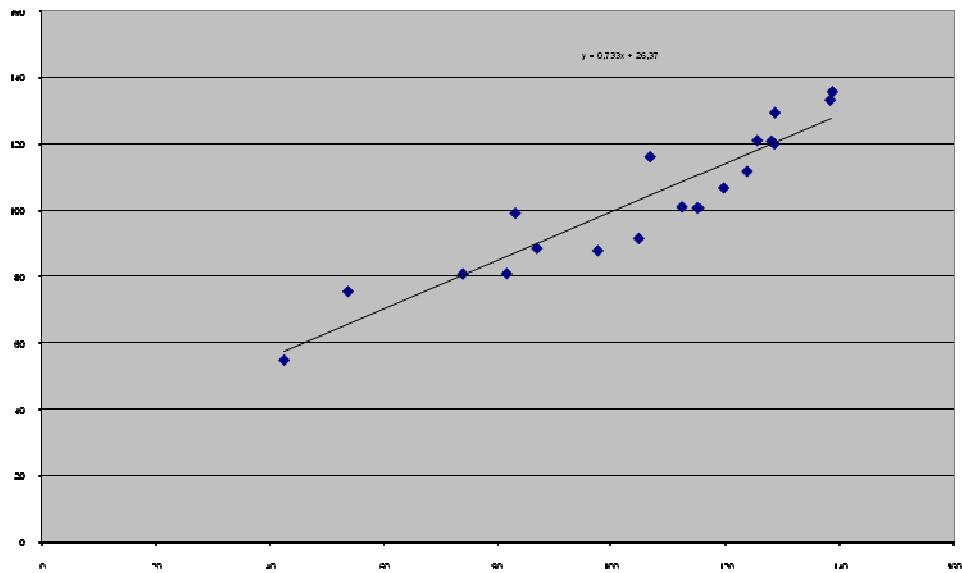


Fig. 2. Retta di regressione per i livelli di occupazione (unità)

Dati per la costruzione delle rette di regressione

Rilevazioni congiunte di:

Tempi di giacenza di singoli container (settimane) e Livelli di occupazione delle due aree (unità)

$$y = 0,4717x + 0,7194$$

$$y = 0,733x + 26,37$$

0,9	1,2
1,9	1,35
1,5	1,46
1,45	1,36
1,1	1,6
1,6	1,2
1,45	1,58
0,94	1,11
1,3	1,05
1,1	1,4
1,9	1,22
1,26	1,48
2,3	1,6
1	1,3
1,52	1,3
1,7	1,44
1,5	1,2
1,49	1,1
2,3	2,4
1,7	1,4
1,71	1,8
1,64	1,37
1,76	2,1
2,25	1,95
1,77	1,43

42,4	54,9
73,8	80,9
83	99,2
106,7	116,2
128,4	129,5
138,5	135,8
138,1	133,3
125,3	121,1
128,3	120,1
127,8	120,9
123,5	111,8
119,4	106,8
115	100,8
112,3	101,1
104,7	91,6
86,8	88,6
97,5	87,9
81,5	81
53,6	75,6

Analisi (puntuale) della correlazione

Si consideri un semplice modello produttore consumatore operante con una logica di tipo “pure push”, cioè con un produttore che immette i prodotti appena realizzati nel buffer in accordo al suo ritmo di produzione e senza tenere conto del ritmo di consumo.

Il sistema è stato generato col metodo Monte Carlo su foglio Excel (tasso di produzione $\lambda = 1$ e tasso di consumo $\mu = 1$, entrambi riferiti a leggi esponenziali) e sono state effettuate alcune stime del coefficiente di correlazione di Pearson per le seguenti coppie di variabili aleatorie d’interesse:

- “tempo fra due produzioni consecutive” (indicata con P);
- “tempo fra due consumi consecutivi”; (indicata C);
- “tempo di giacenza nel buffer”; (indicata con G);
- “tempo di vita (tempo di giacenza + tempo di consumo)” (indicata con V).

Scopo dell’esperimento era appunto quello di verificare che la stima puntuale del coefficiente di Pearson confermasse:

- 1) l’assenza di correlazione fra tempo di produzione e tempo di consumo, vista la logica di funzionamento di tipo “pure push”;
- 2) una forte correlazione positiva fra tempo di vita dei prodotti e tempo di giacenza, segno che il sistema lavora con livelli e tempi di giacenza alti, visto che il ritmo di consumo è solo del 10% più alto del ritmo di produzione;
- 3) una correlazione più o meno significativa, negativa, fra tempi di produzione collocati in un determinato intervallo temporale e tempi di giacenza collocati in un intervallo successivo, atteso che una dilatazione/riduzione dei tempi di produzione consente di diminuire/aumentare le scorte correnti a beneficio/danno dei tempi di giacenza dei prodotti realizzati successivamente.

Per tutte le variabili aleatorie di cui sopra (P, C, G e V) sono state registrate 30 realizzazioni a partire dal prodotto n° 2001 e fino al 2030. Esse sono riportate nella prossima tabella, colonna per colonna. Per la sola variabile P sono state registrate ulteriori 30 osservazioni, riferite al gruppo di prodotti dal n°1971 al 2000, in accordo alla opportunità dell’analisi di correlazione fra tempi di produzione e tempi di giacenza temporalmente sfalsati (punto 3).

Per ricavare la formula della stima puntuale (stima) del coefficiente di correlazione di Pearson si può partire dalla stima (\hat{a}) del parametro “a” (derivata) della retta di regressione e usare una stima per la varianza, dato che risulta:

$$\rho = a \cdot \frac{\sqrt{VAR[X]}}{\sqrt{VAR[Y]}}, \text{ con X e Y variabili aleatorie generiche.}$$

Allora, riprendendo la formula di stima di "a", ricavata in precedenza:

$$\hat{a} = \frac{n \cdot \sum_{i=1}^n x_i y_i - \sum_{i=1}^n x_i \sum_{i=1}^n y_i}{n \cdot \sum_{i=1}^n x_i^2 - \left(\sum_{i=1}^n x_i \right)^2} \quad \text{e riscrivendola così: } \hat{a} = \frac{\left(n \cdot \sum_{i=1}^n x_i y_i - \sum_{i=1}^n x_i \sum_{i=1}^n y_i \right) / n^2}{\left[n \cdot \sum_{i=1}^n x_i^2 - \left(\sum_{i=1}^n x_i \right)^2 \right] / n^2}$$

si perviene alle seguenti (\bar{x} e \bar{y} indicano le medie aritmetiche degli n rispettivi valori)

$$\hat{a} = \frac{\frac{1}{n} \sum_{i=1}^n x_i y_i - \bar{x} \cdot \bar{y}}{\frac{1}{n} \sum_{i=1}^n x_i^2 - \bar{x}^2} = \frac{\frac{1}{n} \sum_{i=1}^n [(x_i - \bar{x}) \cdot (y_i - \bar{y})]}{\frac{1}{n} \sum_{i=1}^n (x_i - \bar{x})^2} \quad (+)$$

e ricordando che risultava pure: $a = \frac{COV(X, Y)}{VAR[X]}$, risulta naturale suggerire il numeratore delle (+) quale formula di stima della covarianza e il denominatore delle (+) quale stima della varianza.

A questo punto ritornando alla: $\rho = a \cdot \frac{\sqrt{VAR[X]}}{\sqrt{VAR[Y]}}$ e disponendo pure della formula di stima della varianza, si ottiene la formula di stima del coefficiente di correlazione di Pearson:

$$\hat{\rho} = \frac{\frac{1}{n} \sum_{i=1}^n x_i y_i - \bar{x} \cdot \bar{y}}{\sqrt{\frac{1}{n} \sum_{i=1}^n x_i^2 - \bar{x}^2} \sqrt{\frac{1}{n} \sum_{i=1}^n y_i^2 - \bar{y}^2}}$$

Applicando questa formula ai dati in tabella si può verificare che risultano i seguenti valori:

$$\hat{\rho}(P, C) = 7.37 \cdot 10^{-5}, \text{ in risposta al punto 1)}$$

$$\hat{\rho}(V, G) = 0.993, \text{ in risposta al punto 2)}$$

$$\hat{\rho}(P_{[1971-2000]}, G) = -0.433 \text{ e } \hat{\rho}(P_{[2001-2030]}, G) = -0.095, \text{ in risposta al punto 3)}$$

Provando a ripetere l'esperimento con tabelle di dati diversi ci si potrà fare un'idea della variabilità della stima puntuale del coefficiente di correlazione di Pearson. Tale variabilità potrebbe generare dubbi sulle osservazioni puntualizzate sopra, in 1), 2) e 3).

P [2001-2030]

C

G

V

P [1971-2000]

0,0799	1,3904	9,4263	10,8167	1,0167
0,1392	0,0699	10,6775	10,7474	0,6407
0,4812	0,4510	10,2661	10,7172	0,2508
0,5034	1,9104	10,2137	12,1242	0,1321
1,8570	2,5478	10,2671	12,8149	0,0812
3,1738	0,0959	9,6411	9,7370	0,3612
1,2192	0,2642	8,5179	8,7820	1,8262
0,0814	0,3477	8,7006	9,0483	0,1758
0,2612	0,0679	8,7871	8,8550	0,1586
0,1636	0,9072	8,6914	9,5986	2,4498
0,0775	0,1815	9,5211	9,7025	1,3183
1,6235	0,0889	8,0790	8,1679	0,6070
0,5616	0,3442	7,6063	7,9505	0,1328
5,0594	0,6896	2,8912	3,5807	0,9086
0,2806	0,1777	3,3001	3,4778	0,6558
0,6495	0,2149	2,8283	3,0431	1,0054
0,3870	0,2807	2,6561	2,9368	3,9550
0,6724	1,3549	2,2645	3,6193	1,0270
1,1667	0,2506	2,4527	2,7033	0,0155
1,0052	0,2268	1,6981	1,9249	4,5030
1,1825	0,1683	0,7423	0,9106	0,7205
0,6949	0,3446	0,2158	0,5604	2,0135
0,8725	0,8444	0,0000	0,8444	0,0056
0,3876	0,6460	0,4568	1,1029	1,1770
0,8284	0,5893	0,2745	0,8638	1,2998
0,7938	1,8298	0,0700	1,8998	0,1909
0,0002	0,7431	1,8996	2,6427	0,8553
0,2010	0,1206	2,4418	2,5624	0,4934
0,1216	0,4480	2,4408	2,8888	2,4960
0,8675	0,1442	2,0213	2,1655	4,9174

Tabella. Campioni di dati ricavati da un’implementazione del modello produttore consumatore col metodo Monte Carlo, (tutti espressi in unità di tempo).

Esercizio sul coefficiente di Pearson

In un laboratorio di informatica è stata allestita una postazione con pc collegato all’intranet universitaria. Ciascuno studente, mediante il proprio pass, ha accesso al laboratorio e può usare detto pc per prenotarsi agli esami aspettando “in fila” il suo turno.

Grazie al meccanismo di accesso al laboratorio mediante pass e al log-in con password al pc, per ciascuno studente è possibile registrare il tempo trascorso all’interno del laboratorio (i.e. tempo di soggiorno) e il tempo che ciascuno ha impiegato per prenotarsi all’esame di interesse (i.e. tempo di servizio che è abbastanza breve).

La considerazione logica del Professore responsabile del laboratorio è che se il tempo di soggiorno è molto simile al tempo di servizio, allora l'attesa in coda per la prenotazione sarà trascurabile e, quindi, un solo pc sarà sufficiente per assolvere a tale servizio. Viceversa, il Professore si aspetta che se il tempo di soggiorno è influenzato da un tempo di attesa significativo (con una coda eccessiva di fronte all'unico pc), allora dovrà essere riscontrata una correlazione debole tra tempo di soggiorno e tempo di servizio.

Per sostenere la plausibilità della tesi del Professore, lo studente è invitato ad usare il coefficiente di Pearson usando i dati registrati, in due momenti diversi, per un campione di cinquanta studenti.

I Rilevamento			
Tempi di Servizio	Tempi di Soggiorno		
2,30	0,04	22,86	13,35
1,52	0,49	23,10	13,27
0,10	0,78	21,66	13,88
0,02	0,00	21,34	12,55
0,15	0,25	20,97	12,20
1,17	0,69	19,68	11,48
0,34	1,94	14,39	12,80
0,97	0,24	13,77	11,71
0,60	0,53	13,57	12,00
0,11	0,90	8,89	11,40
0,04	0,96	8,55	11,90
0,07	0,80	6,52	10,41
0,39	0,38	4,16	10,26
0,44	1,10	4,34	11,24
0,23	0,07	4,40	8,18
2,17	0,72	6,15	8,65
3,72	0,71	9,74	7,47
0,02	0,77	9,57	5,90
0,47	0,27	7,72	4,21
1,88	4,70	9,51	8,52
1,66	0,29	10,54	6,46
2,84	0,24	12,73	6,58
0,12	0,36	12,66	6,83
1,01	0,05	13,41	6,85
0,66	1,60	13,48	4,56

II Rilevamento			
Tempi di Servizio	Tempi di Soggiorno		
0,62	0,01	0,62	0,01
1,09	0,29	1,09	0,29
0,05	1,85	0,05	1,85
1,46	0,84	1,46	2,06
0,80	0,40	0,80	0,40
0,04	1,24	0,04	1,24
0,38	2,39	0,38	2,39
0,16	0,60	0,16	1,51
0,09	1,83	0,09	1,83
0,74	1,14	0,74	1,14
0,82	0,39	0,82	0,39
2,68	0,41	2,68	0,41
0,43	0,20	1,12	0,20
1,77	1,21	1,77	1,21
0,86	0,75	0,86	0,75
0,24	0,68	0,24	0,68
0,07	0,13	0,07	0,13
0,25	0,86	0,25	0,86
1,11	0,02	1,11	0,02
1,40	0,93	1,40	0,93
0,36	0,69	0,36	0,69
0,14	1,17	0,14	1,17
0,66	0,07	0,66	0,07
0,39	0,51	0,39	0,51
1,58	0,90	1,58	0,90

Non conoscendo ancora il procedimento per calcolare il coefficiente di Pearson a partire da realizzazioni indipendenti delle variabili aleatorie X e Y

$$\rho \doteq \frac{COV(X, Y)}{\sqrt{VAR[X] \cdot VAR[Y]}}$$

è stata utilizzata la seguente formula implementata in Excel:

$$\hat{\rho} = \frac{\frac{1}{n} \sum_{i=1}^n x_i y_i - \bar{x} \cdot \bar{y}}{\sqrt{\frac{1}{n} \sum_{i=1}^n x_i^2 - \bar{x}^2} \sqrt{\frac{1}{n} \sum_{i=1}^n y_i^2 - \bar{y}^2}}$$

che sarà studiata più avanti nel corso.

Il coefficiente di Pearson calcolato sul gruppo di dati provenienti da questi due rilevamenti risulta pari a 0,05 (I Rilevamento) e 0,94 (II Rilevamento).

Quali considerazioni si possono fare?