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)
- un evento o risultato “d’interesse” corrisponde, in generale, ad un sottoinsieme dei risultati elementari, cioè alla combinazione logica di eventi elementari;

(quante volte esce testa se lancio la moneta)

- 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 ∈ Ω (***Spazio degli eventi***) si ha P(A) > 0; prob non è nulla

(A2) P(Ω) = 1; per noi tutta la probabilità è pari a 1

(A3) P(A∪B) = P(A) + P(B).

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

(A1’) per ogni evento A ∈ Ω si ha P(A) >= 0;

(A2’) P(Ω) = 1;

(A3’) P(A∪B) = P(A) + P(B) nell'ipotesi che A ∩ B = ∅ (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:

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

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

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

In tal modo, se l'unione all'infinito degli eventi Ai produce un altro evento allora il valore (compreso fra 0 e 1) a cui converge la serie delle P(Ai) 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, …, n, …

E si assegnino le seguenti:

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

Allora:

Serie geometrica!

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

$$= \sum_{i=1}^{\infty} (1-\rho)\rho^{i-1} = (1-\rho)\cdot\sum_{j=0}^{\infty}\rho^{j}$$

La serie geometrica converge a 1/(1-p)

$$= (1-\rho)\cdot\frac{1}{1-\rho} = 1 \qquad \boxed{\text{infatti}: \bigcup_{i=1}^{\infty} A_i \equiv \Omega}$$

8

Continua esempio sull' unione all'infinito di eventi

Si consideri adesso la seguente:

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

Allora:

$$= \sum_{i=n}^{\infty} (1-\rho)\rho^{i-1} = \sum_{j=0}^{\infty} (1-\rho)\cdot\rho^{n-1}\cdot\rho^{j} \qquad (j \triangleq i-n)$$

$$= (1-\rho)\cdot\frac{1}{1-\rho}\cdot\rho^{n-1} = \rho^{n-1} \quad < 1$$

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

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:

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

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) <u>Definizione degli eventi di interesse</u>

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) <u>Calcolo delle probabilità di eventi di interesse</u>

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

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

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

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

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 $E_1 \subset E_2 \subset \ldots \subset E_n \subset E_{n+1} \subset \ldots$. Se, invece, risulta $E_1 \supset E_2 \supset \ldots \supset E_n \supset E_{n+1} \supset \ldots$ allora si dirà che la sequenza è decrescente.

E1 sarebbe superset ovvero insieme che contiene E2.

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

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

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

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

<u>Proposizione</u>:

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:

$$P(E) \triangleq P\left(\lim_{n\to\infty} E_n\right) = \lim_{n\to\infty} P(E_n)$$

si vedrà in futuro cosa vuol dire

<u>PROVA</u> (per una sequenza crescente):

$$P(E) \triangleq P\left(\lim_{n\to\infty} E_n\right) = P\left(\lim_{n\to\infty} \bigcup_{i=1}^{i=n} E_i\right) = P\left(\bigcup_{i=1}^{\infty} E_i\right) = P\left(E_1 \cup \bigcup_{i=1}^{\infty} E_{i+1} \setminus E_i\right)$$

$$= P(E_1) + \sum_{i=1}^{\infty} [P(E_{i+1}) - P(E_i)] = P(E_1) + \lim_{n\to\infty} \sum_{i=1}^{n} [P(E_{i+1}) - P(E_i)] = \lim_{n\to\infty} P(E_n)$$

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, ... e quindi E_1, E_2, ... una sequenza crescente.

Poiché:

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

Allora:

$$P(E) = P\left(\bigcup_{i=1}^{\infty} E_i\right) = P(E_1) + \sum_{i=1}^{\infty} \left[P(E_{i+1}) - P(E_i)\right]$$

$$= (1-\rho) + \sum_{i=1}^{\infty} \left[(1-\rho^{i+1}) - (1-\rho^i)\right]$$

$$= \lim_{n\to\infty} P(E_n) = \lim_{n\to\infty} (1-\rho^n) = 1$$

11

Costruzione di un modello di assegnazione

$$P : [0,1,2,\ldots,n] \to [0,1] \in \Re$$

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

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

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

$$\boxed{P(N = n) = \frac{x^n}{n!} e^{-x}, \quad n = 0, 1, 2, \ldots}$$

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, cioé 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)

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:

  $$\underbrace{n \cdot n \cdots n}_{k\ volte} = n^k$$

  esperimento con risultati ripetibili

  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:

  $$n(n-1)(n-2)\cdots(n-(k-1)) = \frac{n!}{(n-k)!}.$$

  esperimento risultati non ripetibili

Fissando $k \hat{=} n$ (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:

$$\frac{n!}{(n-k)!}=\frac{6!}{(6-2)!}=30$$

se fossero 6 posizioni allora avrei 720 permutazioni possibili

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

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

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)

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

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>=h)

$$\boxed{P = \frac{\binom{k}{h}(n-1)^{k-h}}{n^k} \qquad h = 0,1,\ldots,k}$$

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 <u>a caso</u> e <u>senza rimpiazzo</u> dall'insieme S.

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

$$\binom{n}{k}$$ 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

$$P(r \mid k, d, n) = \frac{\binom{d}{r}\binom{n-d}{k-r}}{\binom{n}{k}}$$

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:

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

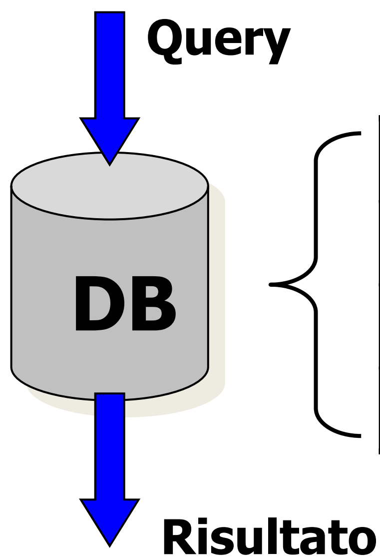
Esempio sulla probabilità ipergeometrica

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:

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



Key	Cognome	…	…	Tel.

Algebra degli eventi

È completamente definita dalle seguenti 5 leggi:

Legge Commutativa: $A \cup B = B \cup A \quad A \cap B = B \cap A$

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

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

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

Legge del Complemento: $A \cup \overline{A} = \Omega \quad A \cap \overline{A} = \varnothing$

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. $A \cup A = A; \quad A \cap A = A; \quad A \cup \Omega = \Omega; \quad A \cap \varnothing = \varnothing;$
2. $A \cap (A \cup B) = A; \quad A \cup (A \cap B) = A;$
3. $\overline{A \cup B} = \overline{A} \cap \overline{B}; \quad \overline{A \cap B} = \overline{A} \cup \overline{B};$
4. $\overline{(\overline{A})} = A; \quad A \cup (\overline{A} \cap B) = A \cup B.$

A titolo di esempio:

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

$(A \cap B) \cup (\overline{A} \cap B)$	
$(B \cap A) \cup (B \cap \overline{A})$	(legge commutativa)
$B \cap (A \cup \overline{A})$	(legge distributiva)
$B \cap \Omega$	(legge del complemento)
B	(legge dell'identità)

Vedremo nelle prossime pagine come:

$P(A \cap B) = P(A) \cdot P(B)$ se e solo se A e B sono indipendenti

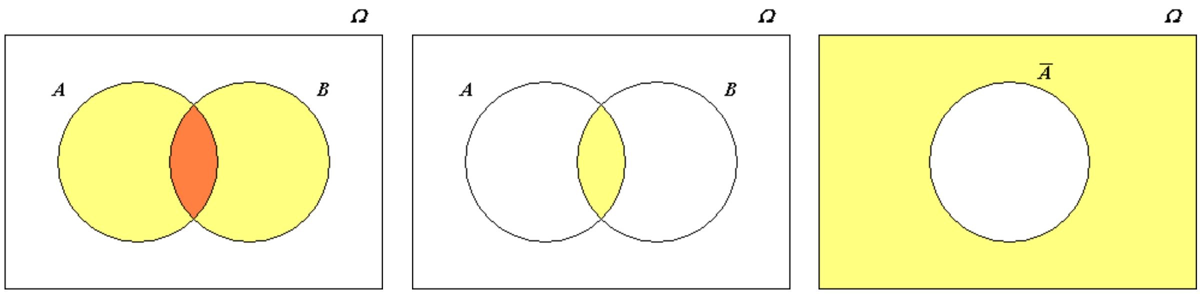
$P(A \cup B) = P(A) + P(B)$ se A e B sono eventi disgiunti (intersezione nulla, NON POSSONO accadere contemporaneamente!)

Esempio evento congiunto: da un sacco di palline colorate e numerate estraggo una pallina rossa e pari (evento congiunto perchè 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 $\overline{A}$. Il diagramma di Venn della seconda figura, invece, mostra che $B \subset A$.



(A) (B) (C)

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

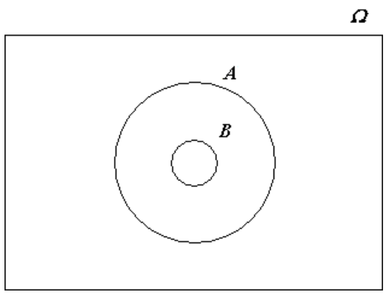
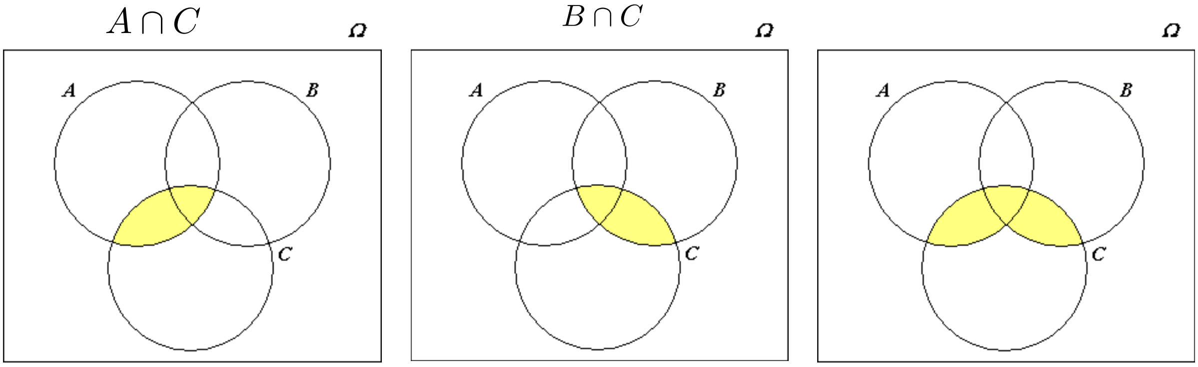


Figura – Diagramma di Venn che illustra la relazione $B \subset A$

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

$A \cap C$ $\qquad\qquad$ $B \cap C$



(A) (B) (C)

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***:

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

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

DEFINIZIONI

- *Gli eventi $A_1, \ldots, A_n$ sono detti **mutuamente esclusivi** se e solo se risulta*

qualsiasi coppia ha intersezione nulla $\quad A_i \cap A_j = \begin{cases} A_i & se \quad i = j \\ \varnothing & altrimenti \end{cases} \quad (i = 1 \ldots n,\ j = 1 \ldots n)$

- *Gli eventi $A_1, \ldots, A_n$ sono detti **collettivamente esaustivi** se e solo se risulta:*

$$A_1 \cup A_2 \cup \ldots \cup A_n = \Omega$$

- *Gli eventi $A_1, \ldots, A_n$ 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

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

(DUE EVENTI)

Prova:

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

$$P(A\cup B)=P(A)+P(\overline{A}\cap B) \qquad P(B)=P(A\cap B)+P(\overline{A}\cap B)$$

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

da cui ottengo che

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

Disuguglianza di Bonferroni

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

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

Disuguaglianza di Boole

(n EVENTI)

probabilità dell'unione di N eventi NON disgiunti

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

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

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

Disuguaglianze di Bonferroni

Nel caso di collezioni finite di eventi, la precedente disuguaglianza 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 $i_1, i_2, \cdots, i_k$ 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:

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

$$\text{Da} \quad B\ certo \implies P(B) = 1$$

Si pone:

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

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

E banalmente si verifica che:

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

Dunque:

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

e si ottiene:

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

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

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

Cattura la dipendenza fra gli eventi.

$$\boxed{\text{Basta porre:} \quad C \triangleq A_2 \cap A_3}$$

$$P\left(A_1 \cap A_2 \cap A_3\right) = P\left(A_1 \cap C\right) = P\left(A_1 \mid C\right) \cdot P(C)$$

come appena visto

esplicito nuovamente

$$= P\left(A_1 \mid C\right) \cdot P\left(A_2 \cap A_3\right) = P\left(A_1 \mid C\right) \cdot P\left(A_2 \mid A_3\right) \cdot P\left(A_3\right)$$

esplicito C

$$= P\left(A_1 \mid A_2 \cap A_3\right) \cdot P\left(A_2 \mid A_3\right) \cdot P\left(A_3\right)$$

$$\boxed{\text{Assumendo:} \quad P\left(A_1 \mid A_2 \cap A_3\right) \triangleq P\left(A_1 \mid A_2\right)}$$

$$= P\left(A_1 \mid A_2\right) \cdot P\left(A_2 \mid A_3\right) \cdot P\left(A_3\right)$$

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

dipendenza a catena

Dipendenza fra eventi generalizzata

È illustrata dalla formula seguente:

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

suggerimento: $B \triangleq A_2 \cap \ldots \cap A_n$

e poi $= P[A_1 \cap B] = P[A_1 \mid B] P[B]$

Formula finale detta Chain Rule:

$$P(A_1 \cap A_2 \cap \ldots \cap A_n) = P[A_1 \mid (A_2 \cap A_3 \cap \ldots \cap A_n)] *$$
$$* P[A_2 \mid (A_3 \cap \ldots \cap A_n)] * \ldots * P[A_{n-1} \mid A_n] * P[A_n]$$

- **Dipendenza a catena generalizzata**

È definita nel seguente modo

$$P(A_1 \cap A_2 \cap \ldots \cap A_n) = P(A_n \cap A_{n-1} \cap \ldots \cap A_1) \triangleq$$
$$\triangleq P[A_n \mid A_{n-1}] * P[A_{n-1} \mid A_{n-2}] *$$
$$\ldots * P[A_2 \mid A_1] * P[A_1]$$

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 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) o P(B)] portandola al valore unitario, ovvero ipotizzare che si verifichi con certezza uno dei due eventi [A 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

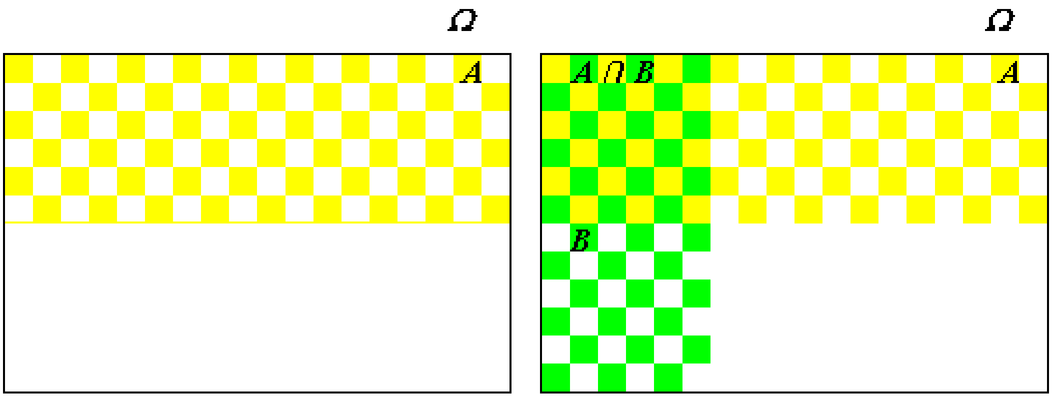
$$P(A \cap B) = P(A|B)P(B) = P(B|A)P(A) = P(A)P(B) = P(B)P(A) > 0$$

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

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:



(A) (B)

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 :

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

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 $\hat{=}$ estrazione di un asso; B $\hat{=}$ 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∩B) è 1/52, come ogni altra carta, perché il mazzo non è truccato. E questo permette di verificare la validità della relazione:

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

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 <u>sufficiente</u> 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: $P(A \cap B \mid B) \equiv P(A \mid B) = 1/13$, 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 ∩ di cuori" (A∩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:

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

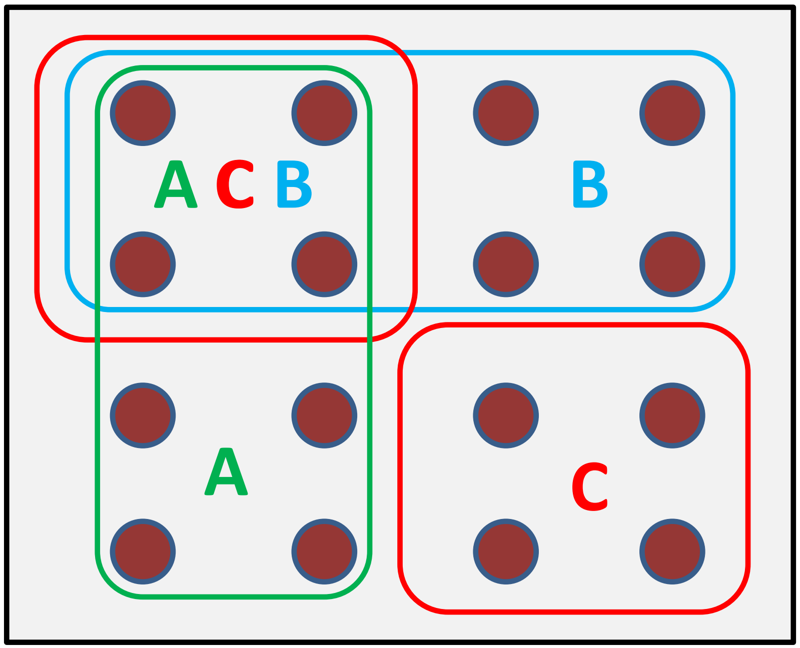
$$+$$

$$\begin{cases} P(A \cap B) = P(A) \cdot P(B) \\ P(A \cap C) = P(A) \cdot P(C) \\ P(B \cap C) = P(B) \cdot P(C) \end{cases}$$

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

Generalizzare al caso di n eventi.

Pairwise independence and mutual independence



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

$$P(A) = P(B) = P(C)$$
$$= 8/16 = 1/2$$

Pairwise independence and mutual independence

P A I R W I S E

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

$$P(A \cap C) = P(A \mid C) \cdot P(C) = P(A) \cdot P(C)$$
$$= (1/2) \cdot (1/2) = 1/4$$

$$P(B \cap C) = P(B \mid C) \cdot P(C) = P(B) \cdot P(C)$$
$$= (1/2) \cdot (1/2) = 1/4$$

NO MUTUAL:

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

Formula della probabilità totale e corollario di Bayes

Data una partizione $\{B_1, B_2, ..., B_n\}$ 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

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

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

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

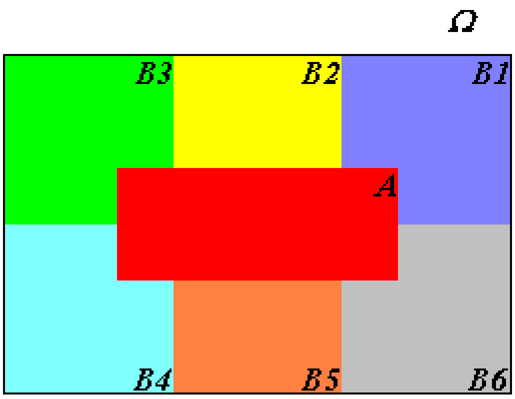


Figura – Illustrazione della formula della probabilità totale per A e $\{B_1, B_2, ..., B_6\}$

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

Corollario di Bayes: Importante

Scelto a caso l'evento Bj della partizione $\{B_1, B_2, ..., B_j, ...B_n\}$ si ricava la formula seguente, nota come regola di Bayes:

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

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

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" $\quad i = M_1, M_2, M_3$
2. A = "pezzo difettoso"

Teorema della PROBABILITÀ TOTALE

$P(A)$ = 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

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

“Conditional Independence” e grafi (di conoscenza)

DEFINIZIONE:

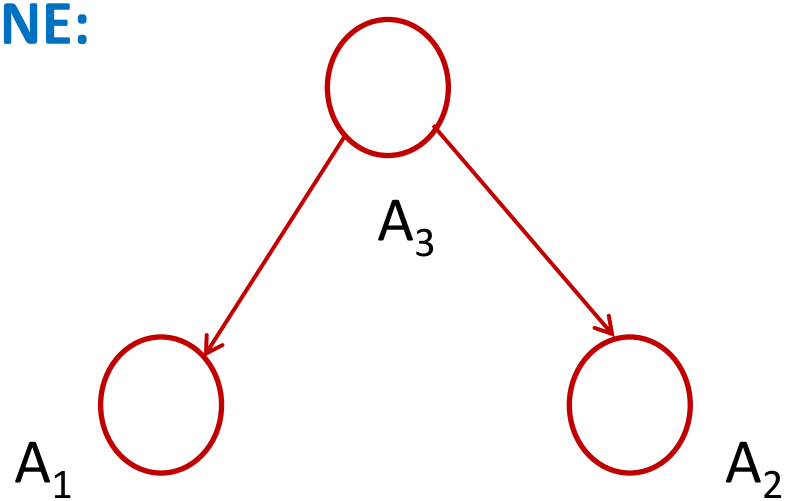
$$P\left(A_1 \cap A_2 \mid A_3\right) \hat{=} P\left(A_1 \mid A_3\right) \cdot P(A_2 \mid A_3)$$

DI CONSEGUENZA:

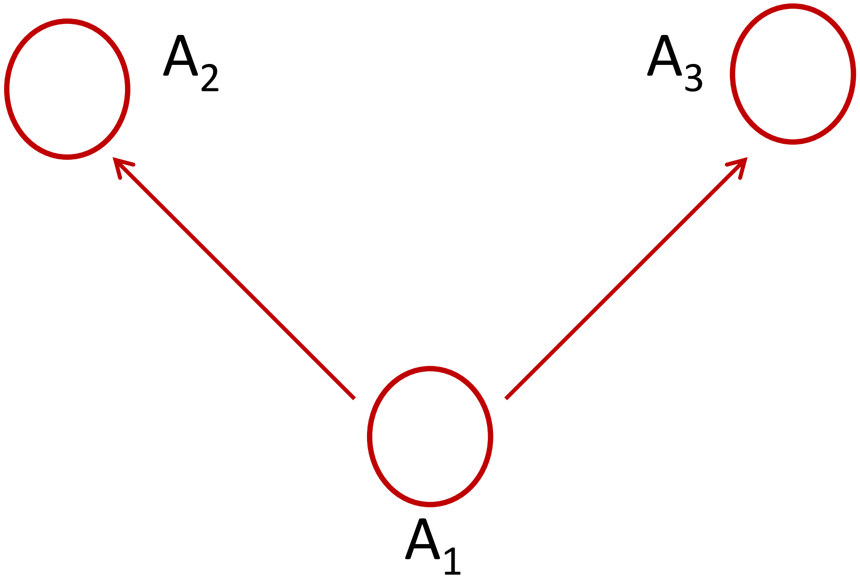
vedi probabilità congiunta e dipendenza

$$P(A_1 \cap A_2 \cap A_3) = P(A_1 \cap A_2 \mid A_3) \cdot P(A_3)$$
$$= P\left(A_1 \mid A_3\right) \cdot P(A_2 \mid A_3) \cdot P(A_3)$$

RAPPRESENTAZIONE:



“Conditional Independence” e grafi



OSSERVAZIONE:

NO arco tra A2 e A3

$$P(A_1 \cap A_2 \cap A_3) = P\left(A_1 \mid A_2 \cap A_3\right) \cdot P(A_2 \cap A_3)$$

$$= P\left(A_1 \mid A_2 \cap A_3\right) \cdot P(A_2) \cdot P(A_3)$$

ALLORA:

$$P(A_2 \cap A_3 \mid A_1) = P\left(A_2 \cap A_3\right) / P(A_1)$$

$$= P(A_2) \cdot P(A_3) / P(A_1) \quad \neq P(A_2) \cdot P(A_3)$$

A2 and A3 are INDEPENDENTS but are NOT CONDITIONAL INDEP

ESEMPIO

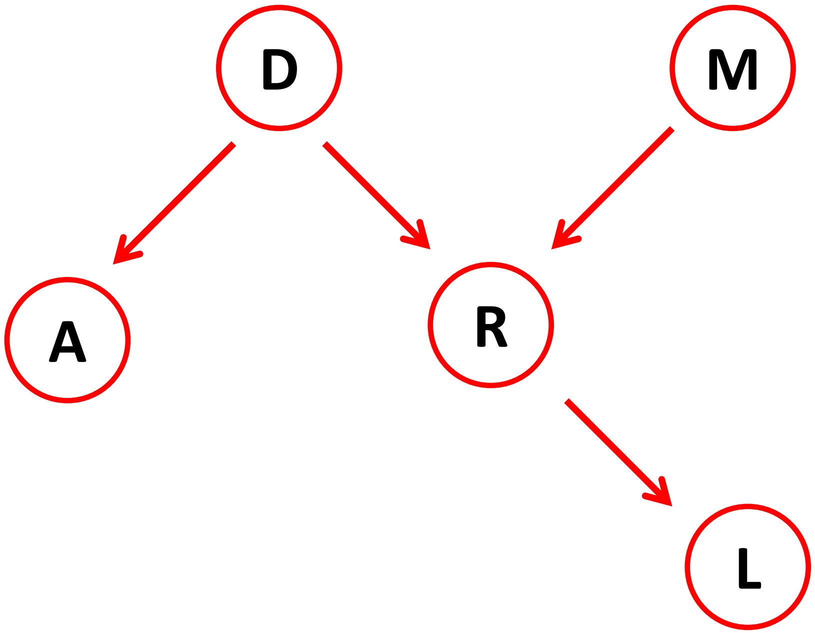
L := "La **L**ezione comincia alle 10 e 50"

R := "Il docente arriva in **R**itardo"

A := "**A**rgomento della lezione è MODELLI su GRAFI."

M := "Forte **M**altempo in atto"

D := "Il **D**ocente 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

$A_i \triangleq$ "componente i-mo funziona", i=1..n,

Affidabilità del sistema seriale

$$R_S \triangleq P\ (\text{"sistema funziona"})$$

$$= P\ (\text{"tutti i componenti funzionano"})$$

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

allora

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

se aggiungiamo l'ipotesi che i componenti sono indipendenti.

Affidabilità del sistema parallelo:

$$\begin{aligned} R_P &= 1 - P(\text{"il sistema non funziona"}) \\ &= 1 - P(\text{"tutti i componenti non funzionano"}) \\ &= 1 - P(\overline{A_1} \cap \overline{A_2} \cap \ldots \cap \overline{A_n}) \\ &= 1 - \prod_{i=1}^{n} P(\overline{A_i}) = 1 - \prod_{i=1}^{n} (1 - P(A_i)) = 1 - \prod_{i=1}^{n} (1 - R_i) \end{aligned}$$

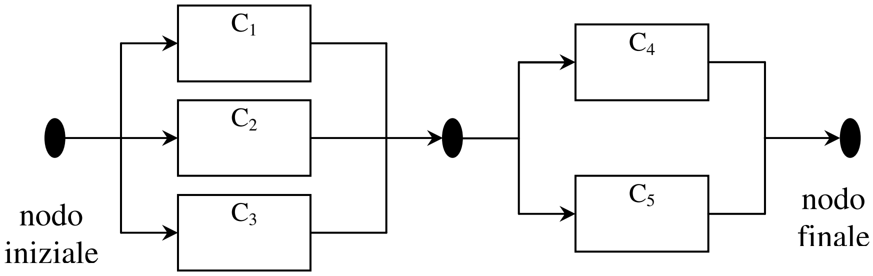
Il sistema smette di funzionare se non vi è alcun, componente funzionante.

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 <u>numerosità</u> e la <u>disposizione</u> dei componenti rilevati funzionanti dall'osservatore casuale sono tali da garantire <u>almeno</u> 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,…,n come eventi disgiunti. Quindi per calcolare l'affidabilità del sistema, occorre applicare la formula della probabilità dell'unione di n eventi non disgiunti:

vista a pagina 16

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

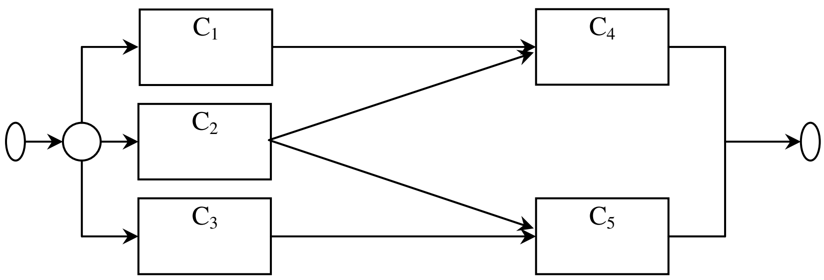
$$= \sum_{\forall i} P(A_i) - \sum_{1 \le i < j \le n} P(A_i \cap A_j) + \sum_{1 \le i < j < k \le n} P(A_i \cap A_j \cap A_k) + \ldots + (-1)^{n-1} P(A_1 \cap A_2 \cap \ldots \cap A_n)]$$

Il metodo dello spazio degli eventi

Quando un sistema è organizzato in un 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

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

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

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:

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

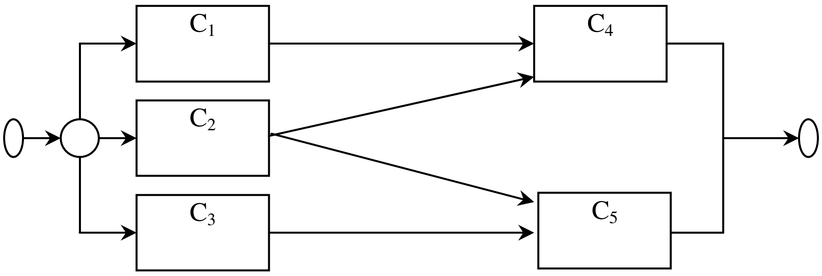
Osservando che i termini dell'unione hanno intersezione non nulla a due a due (A_4 compare nel primo e nel secondo termine, ad esempio), l'affidabilità del sistema può essere calcolata con la formula seguente:

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

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:

$$B \triangleq \text{"}\textit{Componente 2 funziona}\text{"} \quad \text{e} \quad A \triangleq \text{"}\textit{Sistema funziona}\text{"}$$

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

dalla formula della totale

$$P(A) = P(A \cap B) + P(A \cap \overline{B})$$
$$= P(A \mid B)P(B) + P(A \mid \overline{B})P(\overline{B})$$

dopo aver calcolato: $P(A \mid B)$ e $P(A \mid \overline{B})$.

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

Allora: $$P(A \mid B) = 1 - (1 - R_4)(1 - R_5)$$

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!

Allora: $$P(A \mid \overline{B}) = 1 - (1 - R_1 R_4)(1 - R_3 R_5)$$

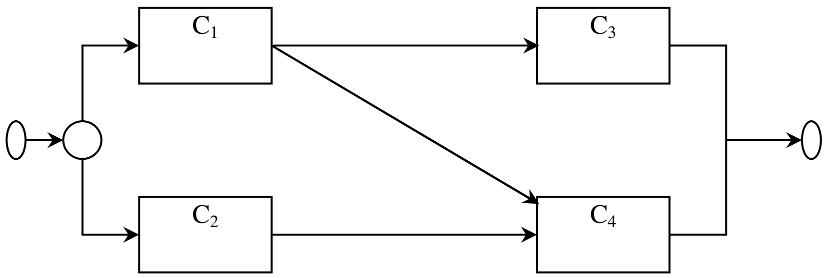
In definitiva:

$$R_{SISTEMA} = [1 - (1 - R_4)(1 - R_5)]R_2 + [1 - (1 - R_1 R_4)(1 - R_3 R_5)](1 - R_2)$$

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 funzioni.

Soluzione

Si definisca l'evento $A \triangleq$ 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,

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

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

$$= (R_{C1} * R_{C3}) + (R_{C2} * R_{C4}) + (R_{C1} * R_{C4}) - (R_{C1} * R_{C3} * R_{C2} * R_{C4}) +$$
$$- (R_{C2} * R_{C1} * R_{C4}) - (R_{C3} * R_{C1} * R_{C4}) + (R_{C3} * R_{C2} * R_{C1} * R_{C4})$$
$$= 0{,}874 + 0{,}833 + 0{,}931 - 0{,}728042 - 0{,}79135 - 0{,}85652 + 0{,}728042$$
$$= 0{,}99013$$

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

$A \triangleq$ il sistema funziona

$B \triangleq$ il componente C_1 funziona

allora

$P(A) = P(A \mid B) * P(B) + P(A \mid \overline{B}) * P(\overline{B}) = P(A \cap B) + P(A \cap \overline{B})$.

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

$P(A \mid \overline{B}) = R_{C2} * R_{C4}$.

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

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

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:

$A \triangleq$ il sistema funziona

$B \triangleq$ il componente C_1 funziona

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

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

Quindi

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

Esempio: Probabilità totale e limite superiore all'Affidabilità

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

$D_i \triangleq$"*Almeno un componente in C_i è guasto*".

In linea di principio si può scrivere:

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

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:

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

PROVA:

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

e riconoscendo che

$P(D_2 \mid \overline{D_1}) =$ Prob "*almeno un componente guasto in C_2 | tutti funzionano in C_1*

risulta

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

Quindi

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

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: $\Omega_n = \{ 2^n \; n-ple \; di \; zero \; e \; uno \}$

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:

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

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

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

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:

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

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:

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

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:

$$\omega \triangleq \left\{ \underbrace{1,1,...,1}_{n_1 volte}; \underbrace{2,2,...,2}_{n_2 volte};...; \underbrace{k,k,...,k}_{n_k volte} \right\} \text{ e } \quad (n_1 + n_2 + ... + n_k = n)$$

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

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 indipendete!

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 = "n° tentativi per realizzare k successi". Con k fissato a priori e n° tentativi (n) libero, purché sia n>=k
- B = "esattamente *k-1* successi in *n-1* tentativi" successi fissati quindi prove di Bernoulli!
- C = "il tentativo *n*-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:

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

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:

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

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.

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

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

Cioè il sistema funziona quando almeno “m” componenti su “n” funzionano. Cioè evento A(>=m)

$$R_{m|n} = P(\text{“}\textit{almeno m componenti funzionano}\text{”}) = P(A(>= m))$$

Partizionando lo spazio dei risultati Ω si può scrivere

$$R_{m|n} = P(\bigcup_{i=m}^{n}(\text{ ”}\textit{esattamente i componenti funzionano}\text{”}))$$

$$= \sum_{i=m}^{n} P((\text{”}\textit{esattamente i componenti funzionano}\text{”}))$$

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

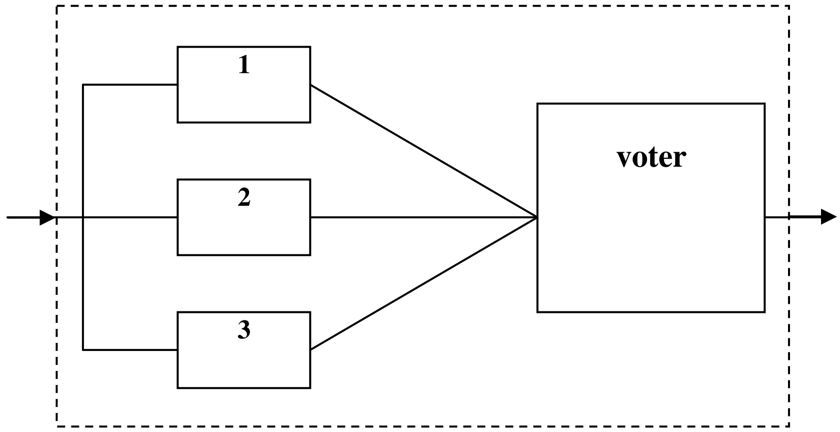
$$R_{m|n} = \sum_{i=m}^{n}\binom{n}{i} R^{i}(1-R)^{n-i}$$

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 disacccordo 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 due 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:

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

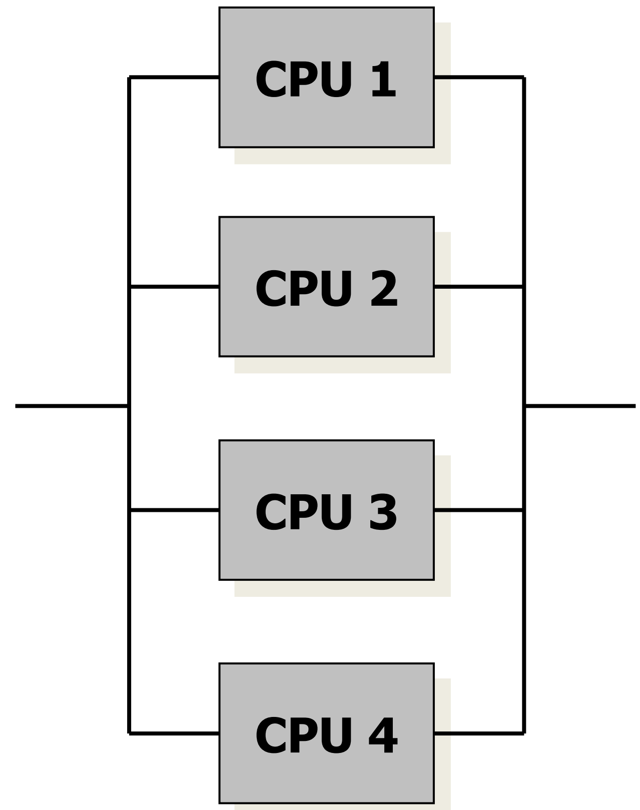
dunque sostituendo n=3 ed m=2

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

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 funzioni?
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 funzioni 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:

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

Perciò:

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

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:

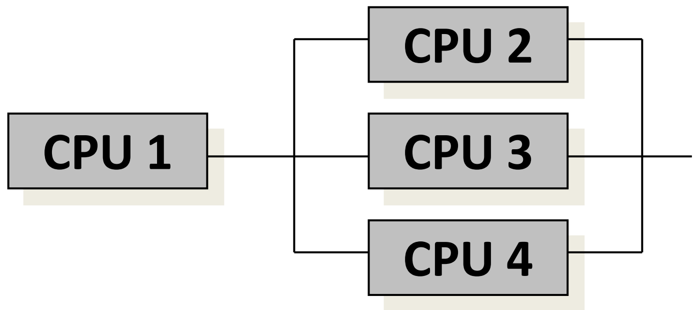
$$P(E) = \frac{P(F \cap C_2)}{P(F)} = \frac{0.0486}{0.9963} = 0.0487$$

<u>Osservazione</u>:
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:

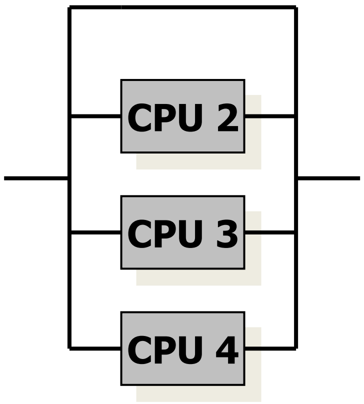


$$R = R_{CPU1} \cdot R_{1|3} = R_{CPU1} \cdot \left(1 - (1 - R_2)(1 - R_3)(1 - R_4)\right)$$

$$= 0.9 \cdot \left(1 - 0.9)(1 - 0.9)(1 - 0.9)\right) = 0.8991$$

Quesito 4:

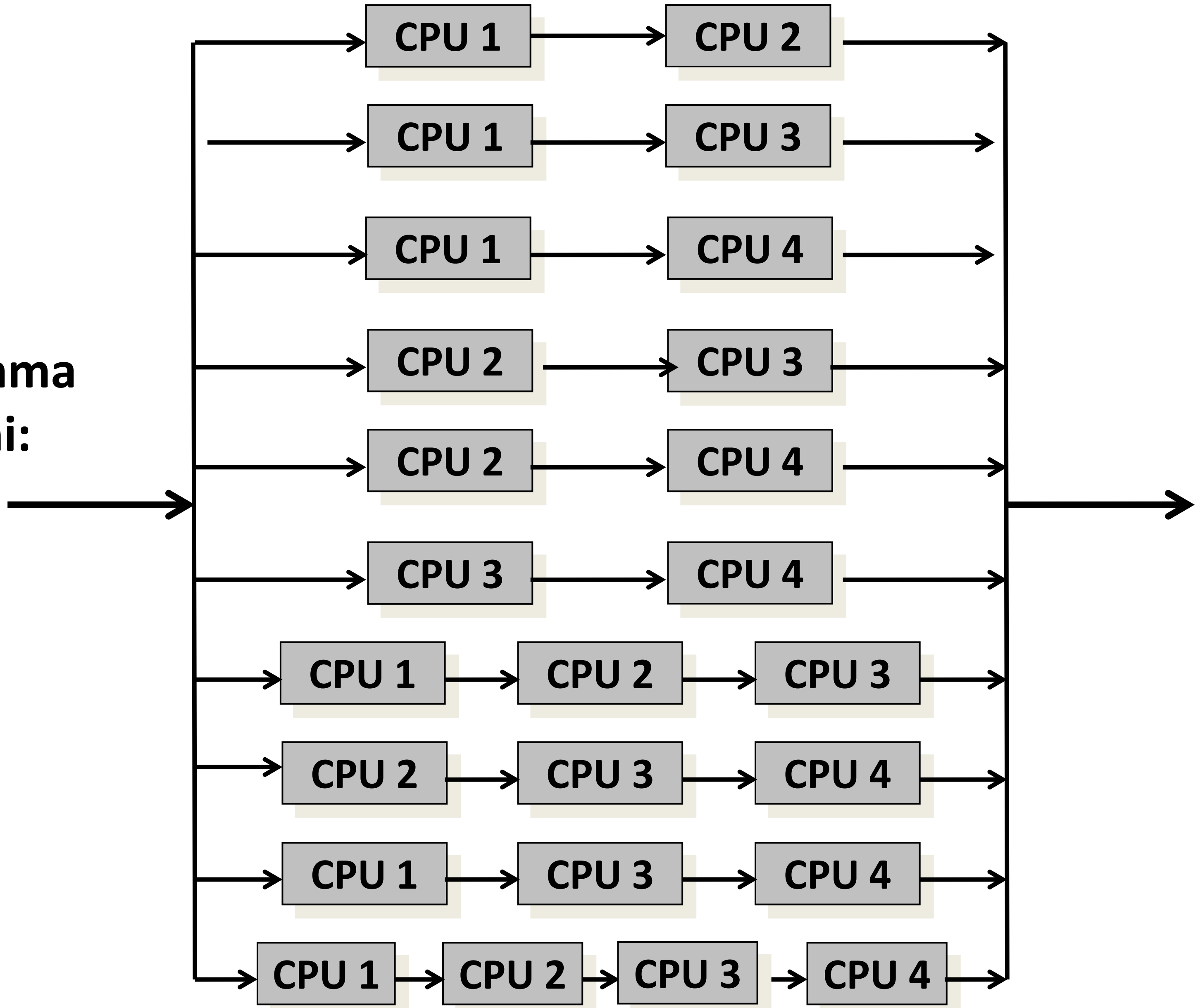
R = P(2-out-of-4 |CPU 1 funziona certamente) = P(1-out-of-3)



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

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,...4)» di processori richieda di accedere al bus nello stesso ciclo di clock è nota, per ipotesi, dalla tabella seguente:

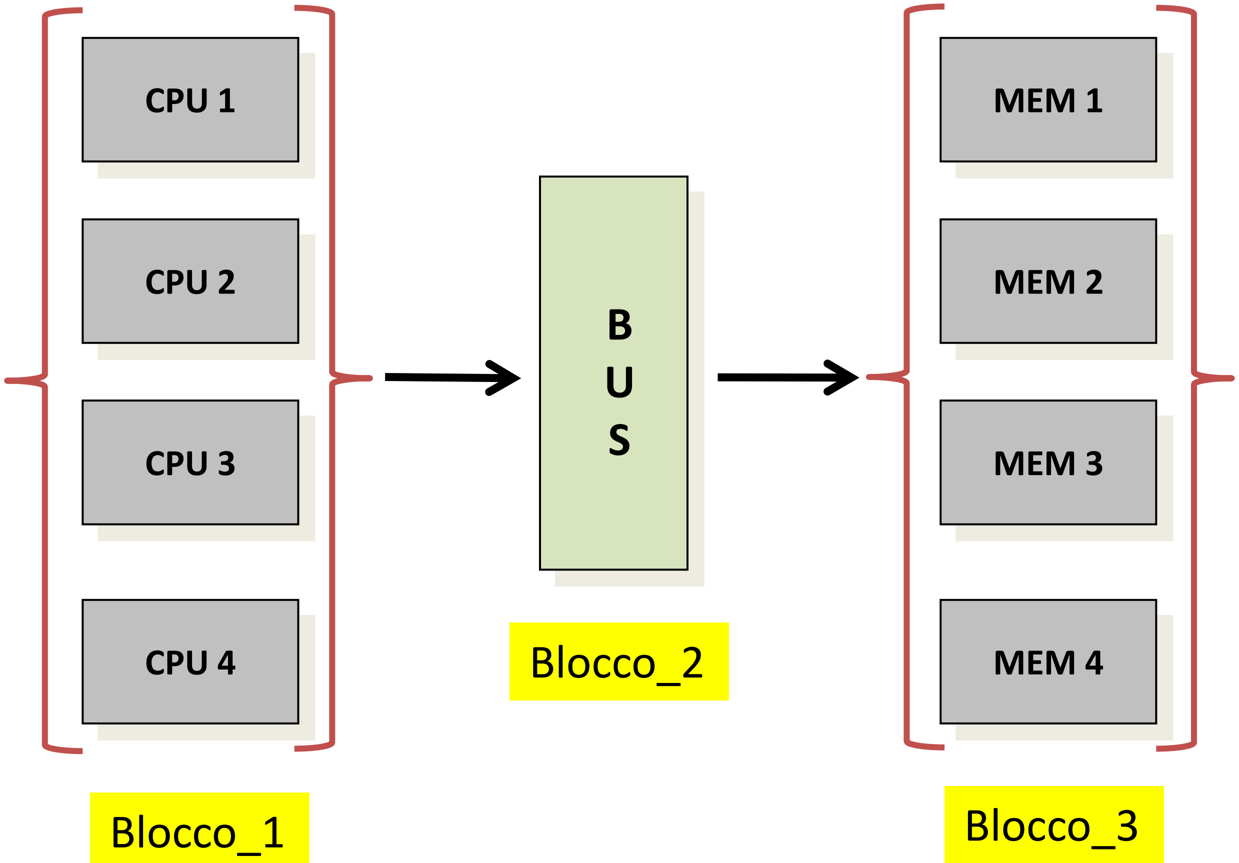
n° proc. richied.	0	1	2	3	4
probabilità	0,2	0,40	0,20	0,12	0,08

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 occorrano 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



<u>IPOTESI</u>:
Il sistema funziona quando funzionano almeno 1 CPU, il BUS e almeno 2 MEM

<u>Modello risultante</u>:

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

Esercizio 1: risposta al quesito 1.

- Sia B_n l'evento: «n (n=0,1,…,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:

$$P(E) = P\left(\bigcup_{n=2}^{n=4} B_n\right) = \sum_{n=2}^{n=4} P(B_n) = 0{,}20 + 0.12 + 0.08 = 0.40$$

<u>Osservazioni</u>:

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

Gli eventi [Bn, n=0,1,…,4] avrebbero potuto non essere elementari, ma calcolabili a partire da altri eventi elementari, cioè direttamente osservabili, ad es. [Ci, i=1,…,4], con Ci =«il processore i-esimo richiede il bus».

In tal caso, sarebbe sorta una difficoltà: P(Ci|Cj,Ck,Cl) =? …

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, …, 4, occorre calcolare le seguenti:

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

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

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

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

$$P(A) = \sum_{n=1}^{n=4} P(A \mid \hat{B}_n) \cdot P(\hat{B}_n) =$$

$$= 1 \cdot 0.50 + 0.5 \cdot 0.25 + 0.33 \cdot 0.15 + 0.25 \cdot 0.10 = 0.70$$

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»;

<u>Osservazione</u>:

- $P(E_3)$ 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:

$$P(F) = 1 - [P(E_1)) + P(E_2) + P(E_3)]$$
$$= 1 - [0.7 + 0.21 + 0.063] = 1 - 0.973 = 0.027$$

(potendo ripetere la richiesta all'infinito)

<u>Osservazione</u>: 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: $$P(G_{k|n}) = P(G_{k-1|n-1} \mid S) \cdot P(S), \quad \text{con} \quad P(S) = P(A)$$

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

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

altro non è che la dimostrazione della binomiale negativa

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

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|Dbi", i=1,…,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:

$$P(A) = P(A/Db1)*P(Db1) + P(A/Db2)*P(Db2) + \\ +P(A/Db3)*P(Db3) + P(A/Db4)*P(Db4) = 0.049$$

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”

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

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

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

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

<u>Osservazione</u>: 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.

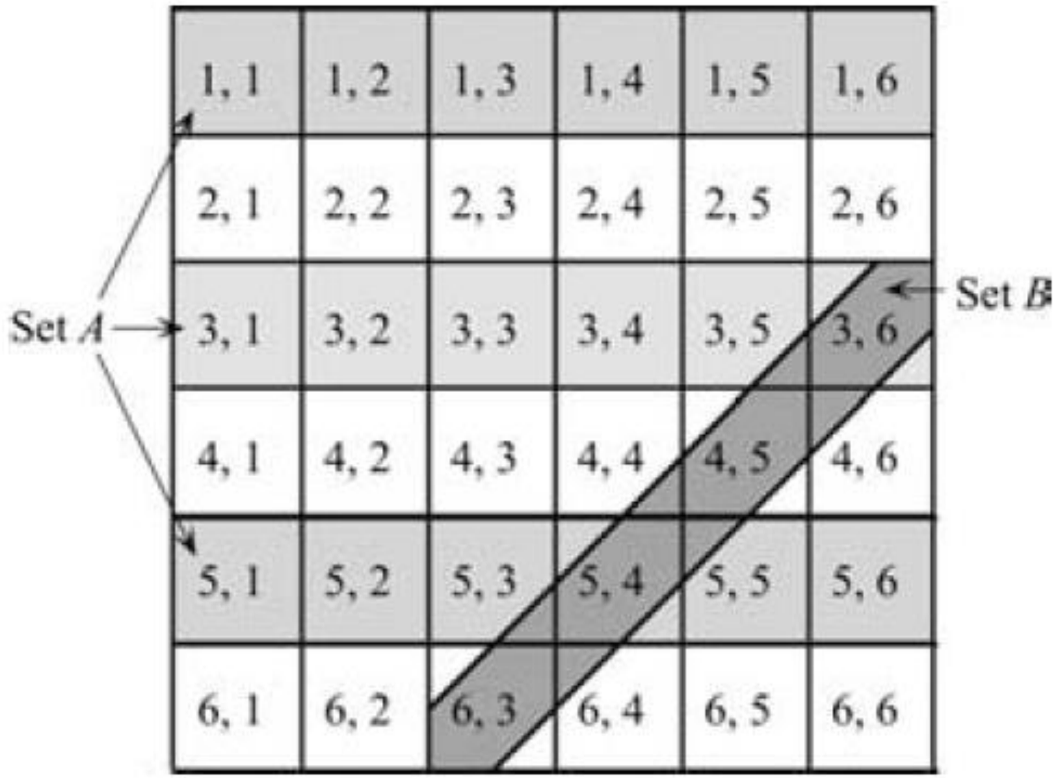

1, 1
1, 2
1, 3
1, 4
1, 5
1, 6
2, 1
2, 2
2, 3
2, 4
2, 5
2, 6
3, 1
3, 2
3, 3
3, 4
3, 5
3, 6
4, 1
4, 2
4, 3
4, 4
4, 5
4, 6
5, 1
5, 2
5, 3
5, 4
5, 5
5, 6
6, 1
6, 2
6, 3
6, 4
6, 5
6, 6
Set A
Set B

FIGURE 2.3.1

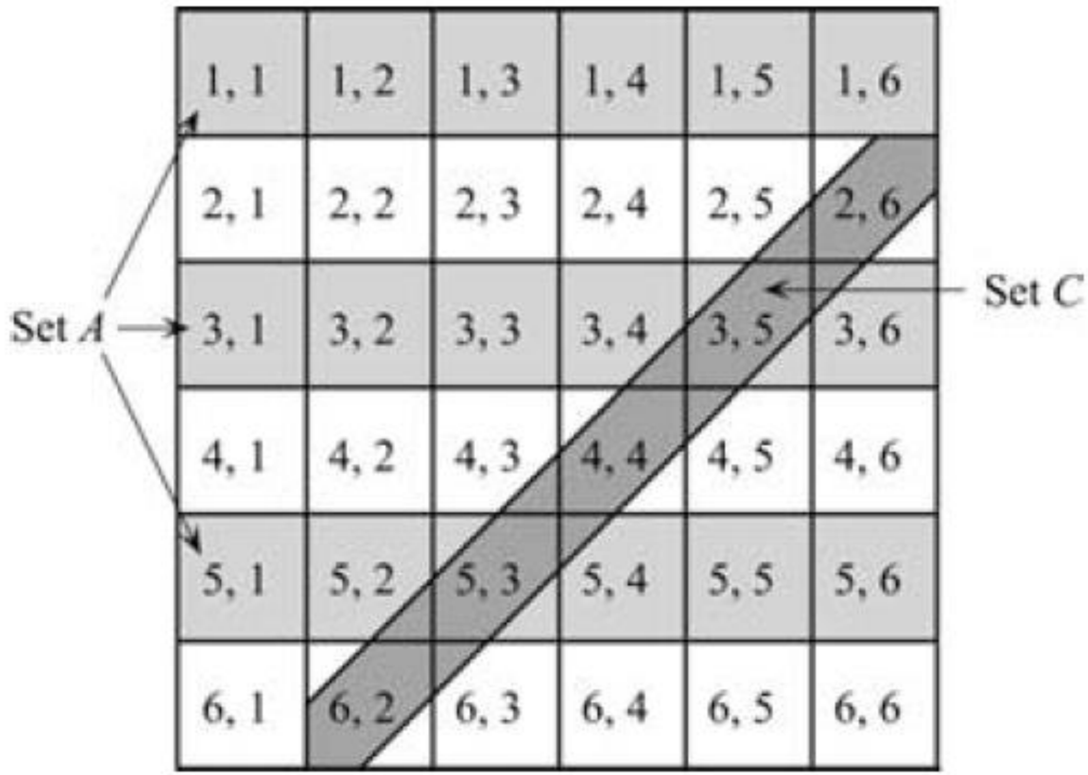

1, 1
1, 2
1, 3
1, 4
1, 5
1, 6
2, 1
2, 2
2, 3
2, 4
2, 5
2, 6
3, 1
3, 2
3, 3
3, 4
3, 5
3, 6
4, 1
4, 2
4, 3
4, 4
4, 5
4, 6
5, 1
5, 2
5, 3
5, 4
5, 5
5, 6
6, 1
6, 2
6, 3
6, 4
6, 5
6, 6
Set A
Set C

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 $P\{B \mid A\}$ 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 $\{B \mid A\} = \{(3,6), (5,4)\}$. Hence $P\{B \mid A\} = \frac{2}{18} = \frac{1}{9} = P\{B\}$, 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

Die A \ Die B	1	2	3	4	5	6
1	2	3	4	5	6	7
2	3	4	5	6	7	8
3	4	5	6	7	8	9
4	5	6	7	8	9	10
5	6	7	8	9	10	11
6	7	8	9	10	11	12

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

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

as an infinite geometric series:

P(*i* in the first row AND win) $= \dfrac{p^2}{p+\frac{1}{6}}$ 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

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

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:

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

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

$$P\{B\} = P\{B \mid A\}P\{A\} + P\{B \mid B\}P\{B\} + P\{B \mid C\}P\{C\}$$
$$= \frac{1}{2} \cdot \frac{1}{3} + 0 \cdot \frac{1}{3} + 1 \cdot \frac{1}{3} = \frac{1}{2}$$

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:

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

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

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	$P\{B\} = 1 \cdot \frac{1}{3} = \frac{1}{3}$

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: $X : \Omega \rightarrow \Re$ Processo stocastico: $X_\omega : \Re \rightarrow S$ con $\omega \in \Omega, S \subseteq \Re$

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) $\rightarrow$ variabili aleatorie discrete (a valori nel discreto).

Caso 2) $\rightarrow$ 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 $P:[0,1,2,...,n] \to [0,1] \in \Re$ che ripartisce i singoli valori di probabilità alle singole realizzazioni, k=0,1,…,n, della K rispettando gli assiomi della probabilità.

$$P(K=k \mid n) \equiv P_K(k \mid n) \triangleq \binom{n}{k} p^k q^{n-k}, \qquad k=0,1,\ldots,n\,;$$

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

1) $0 \le \binom{n}{k} p^k q^{n-k} \le 1 \qquad \forall\, k=0,1,\ldots,n$ dunque probabilità compresa tra 0 e 1

2) $\sum_{k=0}^{n} \binom{n}{k} p^k q^{n-k} = (p+q)^n = 1$ 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:

$$P(S=k \mid \infty) \equiv P_S(k \mid \infty) \triangleq q^{k-1} p, \qquad k=1,\ldots,n,\ldots$$

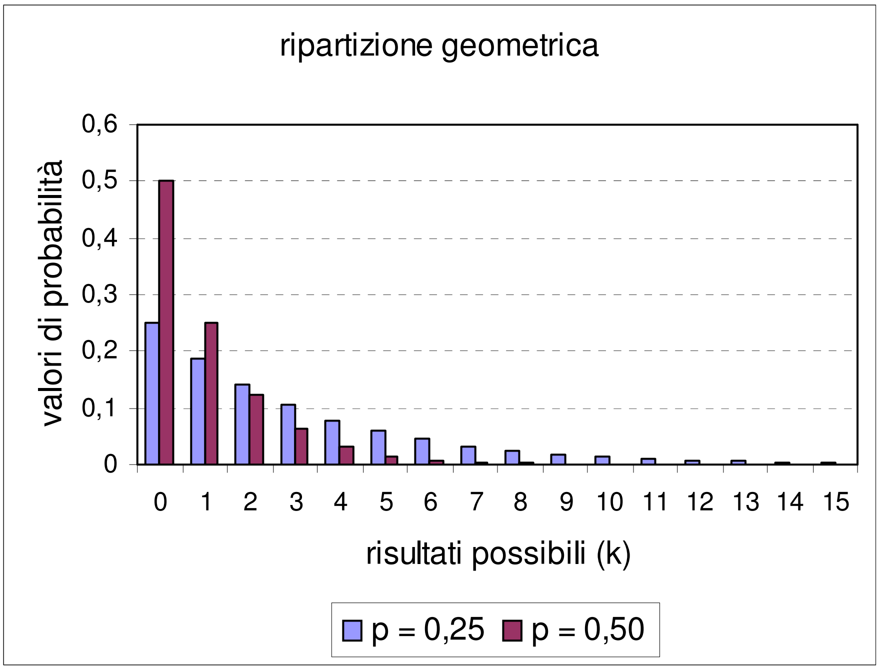
si parte da 1 perché deve accadere sto successo

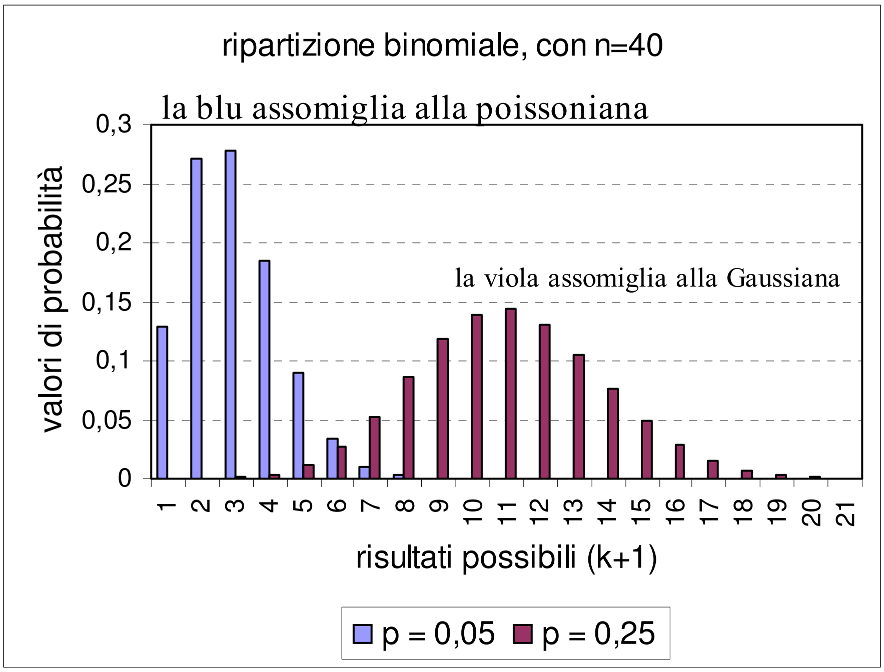
e si verifica che:

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

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)

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

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

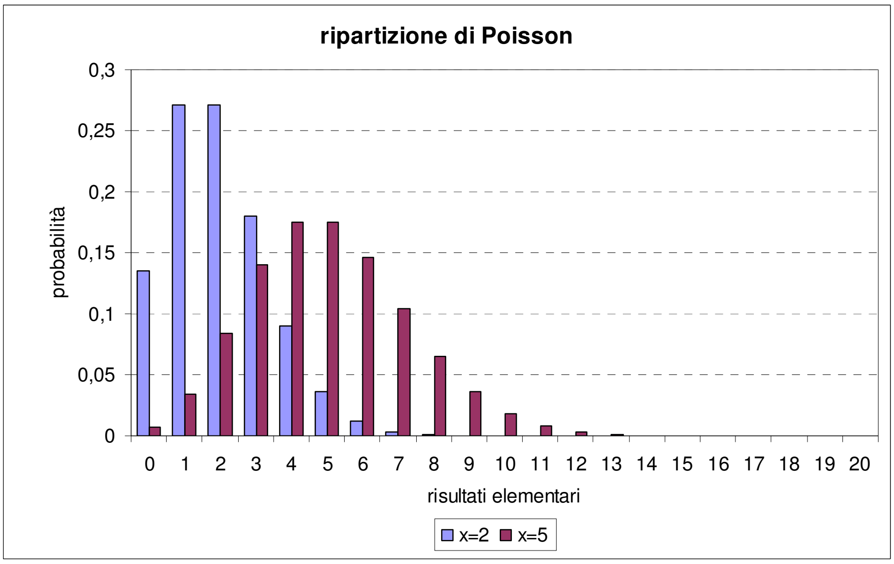
$$\sum_{n=0}^{\infty} \frac{x^n}{n!} e^{-x} = 1 \qquad x > 0$$

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 $n = 0,1,2,...$ 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,…,n,…, si potrebbe adottare la legge geometrica per ripartire valori di probabilità alle possibili realizzazioni:

$$P(K = k \mid \infty) \equiv P_K(k \mid \infty) \triangleq q^k p, \qquad k = 0,1,\dots,n,\dots$$

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:

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

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:

$$0 \le C(m) \le 1 \quad \forall m \qquad \text{e} \qquad C(m) \to 1 \quad se \;\; m \to \infty$$

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:

$$F_X : [0 \le x < \infty] \to [0,1] \in \Re$$

codominio da 0 a 1 perché è una probabilità

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

$$P(X \le x) \equiv F_X(x) \hat{=} \int_{t=0}^{x} f_X(t)dt$$

l'estremo superiore è x, dunque F_X è una funzione integrale

l'integrale sarebbe la somma delle probabilità

Proprietà di F_X

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

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

$$F_X(x_2) \ge F_X(x_1), \quad x_2 \ge x_1$$ è monotona

se x=0 allora la funzione è pari a 0

"alla fine hai terminato"

$$F_X(x \to 0) \to 0, \quad F_X(x \to \infty) \to 1$$

ovvero probabilità è pari 1

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 $f_X : [0 \le x < \infty] \to [0, M] \in \Re^+, \quad M > 1$ 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 finisca tra 30s e 31s. Quindi dx =1s.

Una densità per x che tende a infinito deve fare 0

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

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

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

data la probabilità tra x1 e x2

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

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

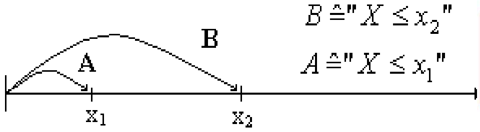
$$\cong f_X(x)\Delta x, \quad \Rightarrow P(X = x) = 0$$

risulta nulla la probabilità associabile ad un punto

Punti di discontinuità nella F_X

Si farà riferimento a due risultati, "$X \le x_1$" e "$X \le x_2$" con $x_1 \le x_2$, 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:



$B \triangleq "X \le x_2"$

$A \triangleq "X \le x_1"$

$$\boxed{\Rightarrow \overline{A} \cap B \equiv (x_1, x_2]}$$

Partendo dalle seguenti relazioni:

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

si riconosce che:

$$P(\overline{A} \cap B) \equiv P(x_1 < X \le x_2)$$

mentre

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

quindi

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

A questo punto, il fatto che $x_2 \to x_1$ può essere visto in termini dell'evento $(\overline{A} \cap B)$ - 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 definisce il modello matematico 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:

$$Pr\{X \le t + \Delta t \mid X \ge t\} \triangleq Pr\{X \le \Delta t\} \qquad (*)$$

probabilità che l'evento finisca prima di t+Δt sapendo che arriva a t sicuramente è pari alla probabilità che finisca 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

$$Pr\{Y \le \Delta t \mid X \ge t\} = Pr\{X \le \Delta t\}$$

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

$$Pr\{t \le X \le t + \Delta t\} = Pr\{X \le t + \Delta t \mid X \ge t\} Pr\{X \ge t\}$$

il complemento al minore uguale è maggiore uguale per questo ho uguale

imponendo la (*)

$$= Pr\{X \le \Delta t\} Pr\{X \ge t\}$$

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 $\lim_{\Delta t \to 0}$ si ottiene una semplice equazione differenziale:

$$\frac{d}{dt}F(t) = \frac{d}{dt}F(t)\Big|_{t=0}[1 - F(t)]$$

è un'equazione differenziale

che ammette come soluzione appunto la <u>legge esponenziale</u>:

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

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

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.

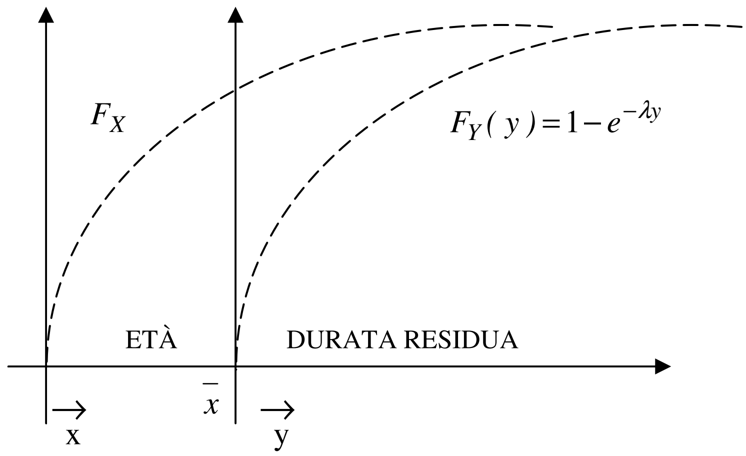
Il ragionamento appena sviluppato porta a definire la *distribuzione esponenziale*:

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

con la seguente *densità esponenziale* :

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

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

Prova: applicando congiunta e condizionta

$$P(Y \leq y \mid X \geq \bar{x}) = P(X - \bar{x} \leq y \mid X \geq \bar{x}) = \frac{P(\bar{x} \leq X \leq \bar{x}+y)}{P(X \geq \bar{x})} = \boxed{\frac{F(\bar{x}+y)-F(\bar{x})}{1-F(\bar{x})}} =$$

$$= \frac{1-\exp\{-\lambda(\bar{x}+y\}-(1-\exp\{-\lambda\bar{x}\})}{\exp\{-\lambda\bar{x}\}}$$

$$= 1 - exp\{-\lambda y\} = F_X(y) \triangleq P(X \leq y)$$

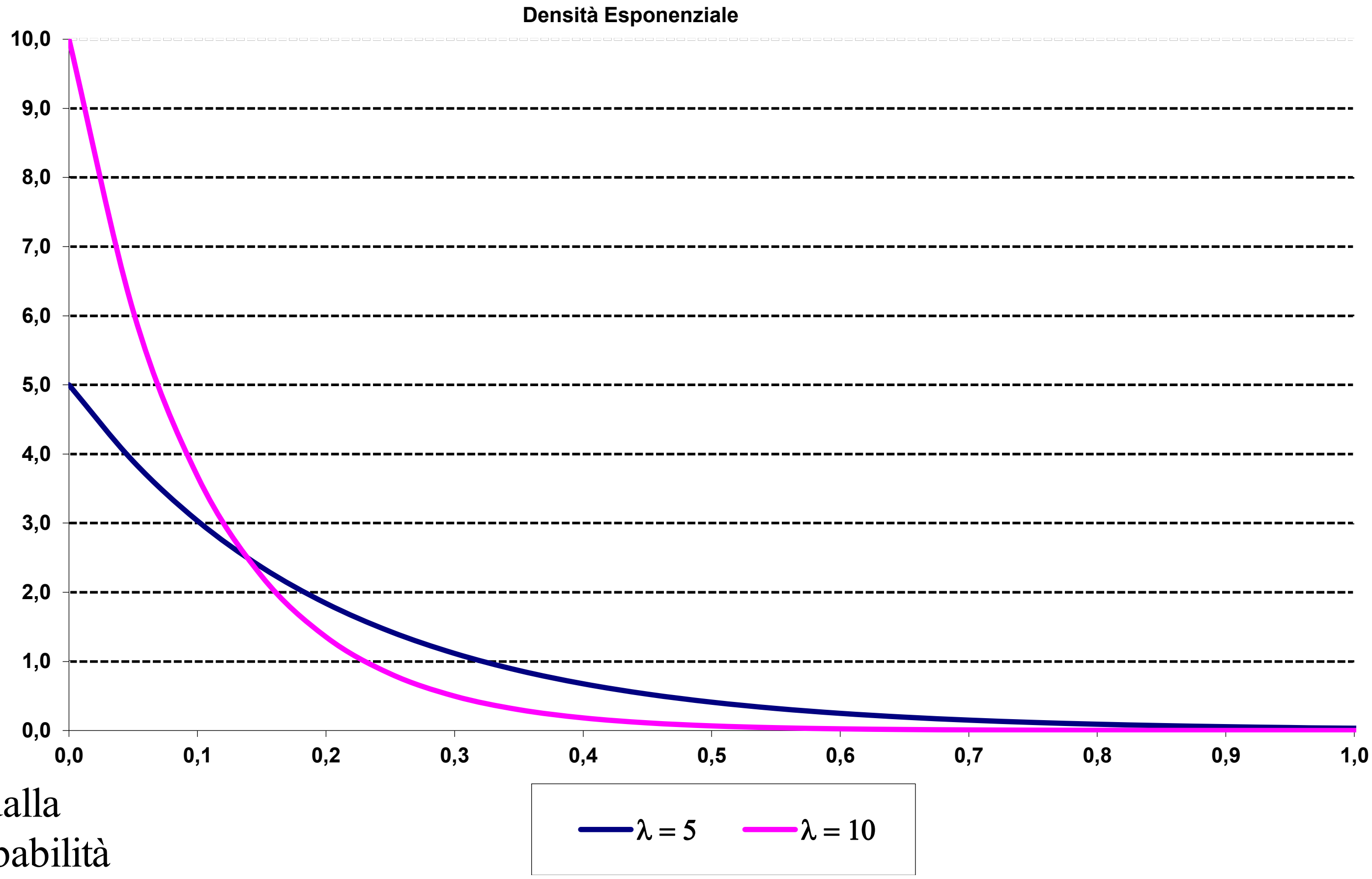
La funzione di distribuzione della vita residua coincide con quella della vita intera, però traslata di una quantità pari all'età.

La funzione densità esponenziale

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

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

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



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 $F_X(t)=1-e^{-\lambda \cdot t}$ 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:

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

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

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

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):

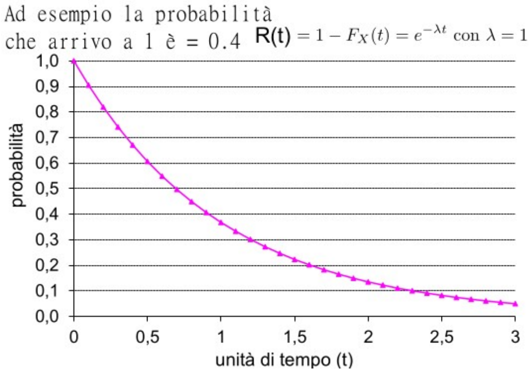
$$R(t) \triangleq P(X \geq t) = 1 - P(X \leq t) = 1 - F_X(t)$$

F(t) è una cumulativa, dunque R(t) è una cumulativa

<= perché continua

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



Per ipotesi, dunque, deve risultare:

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

Proprietà caratteristica è invece:

$$R(t \to \infty) \to 0$$

Nota bene: l'assenza di memoria in affidabilità significa assenza di usura!

Probabilità del guasto

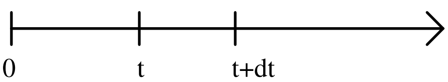
Faccio la derivata di R(t) così vedo cosa accade in un punto, dunque da

$$\frac{d}{dt}R(t) = \frac{d}{dt}(1 - F_X(t)) = -f_X(t)$$

(essendo l'affidabilità decrescente mi aspetto il meno nel risultato della derivata ricordalo)

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 <u>tasso di guasto</u>.

A differenza della funzione affidabilità, il tasso di guasto è una misura <u>istantanea</u> 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 un istante di tempo futuro che si raggiunge si chiama età

$$h(t) \doteq \lim_{\Delta t \to 0} \frac{\text{Prob}(t < X \leq t + \Delta t \mid X > t)}{\Delta t}$$

ignorando delta t e sostituendo con congiunta e condizionante

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

rapporto incrementale ovvero $f_X(t)$

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

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 $h(t) = \lim_{\Delta t \to 0} \frac{R(t) - R(t + \Delta t)}{\Delta t R(t)} = \frac{-R'(t)}{R(t)}$ che messa assieme a $f_X(t)dt = -dR(t)$ porta a $R(t) \cdot h(t)dt = f_x(t)dt$

h(t) sarebbe una velocità.
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

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

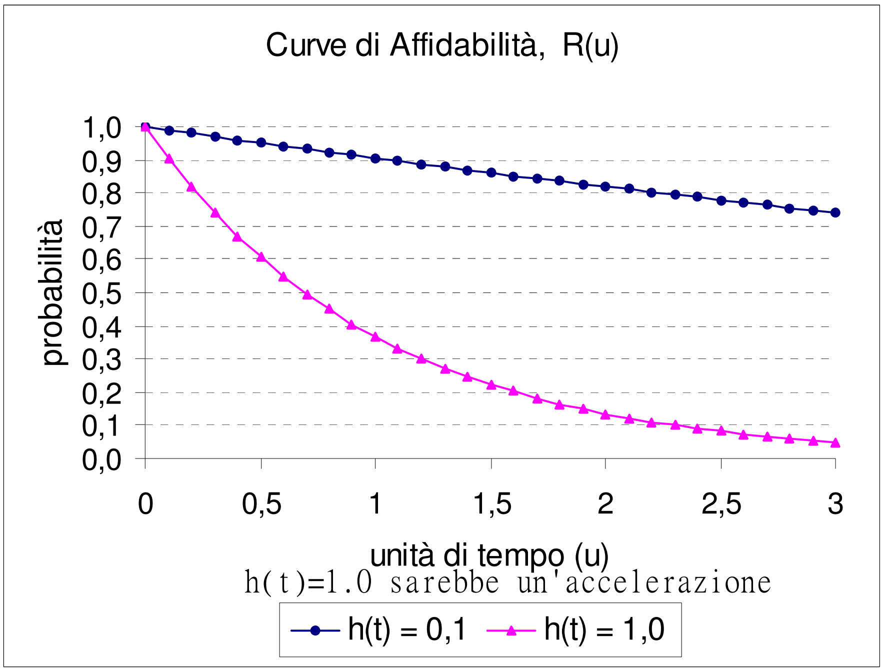
Caso particolare: $h(t) = \lambda \implies R(u) = exp\{-\lambda \cdot u\}$

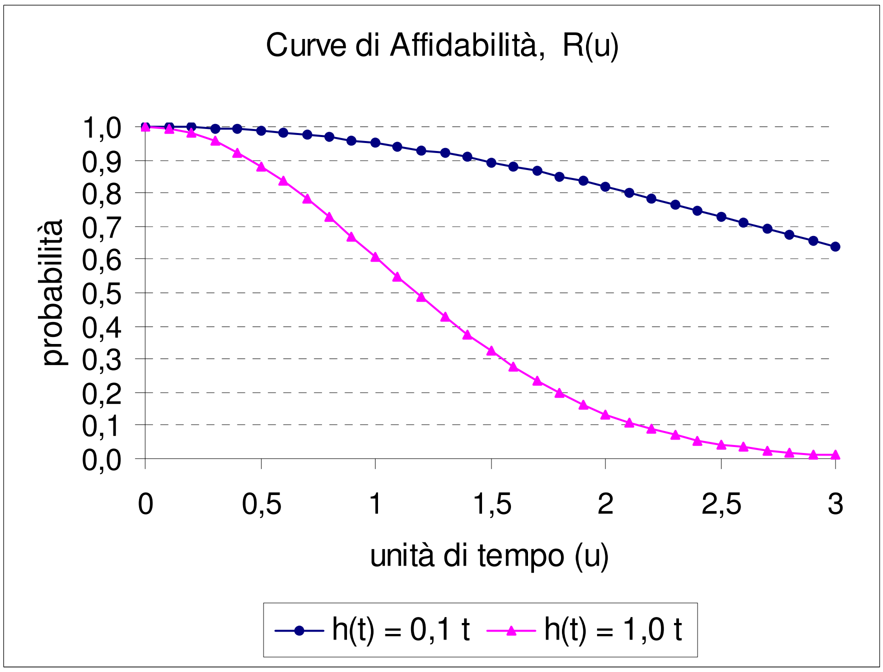
La funzione h(t) è il "modello di guasto" alla base della funzione affidabilità.

(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

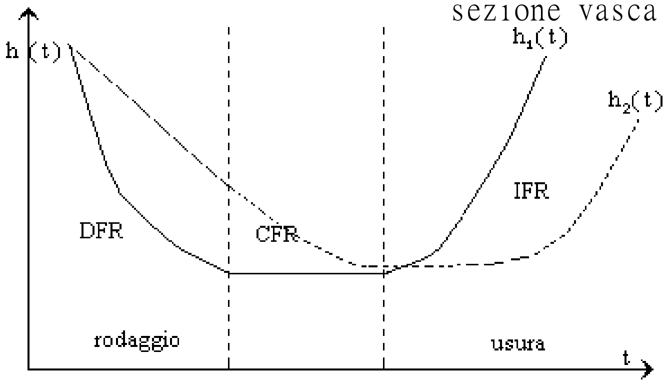




Modelli di guasto

La curva è detta sezione vasca da bagno

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 : $h(t) \triangleq \lambda \alpha t^{\alpha-1}, \quad \lambda > 0, \alpha > 0$.

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

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

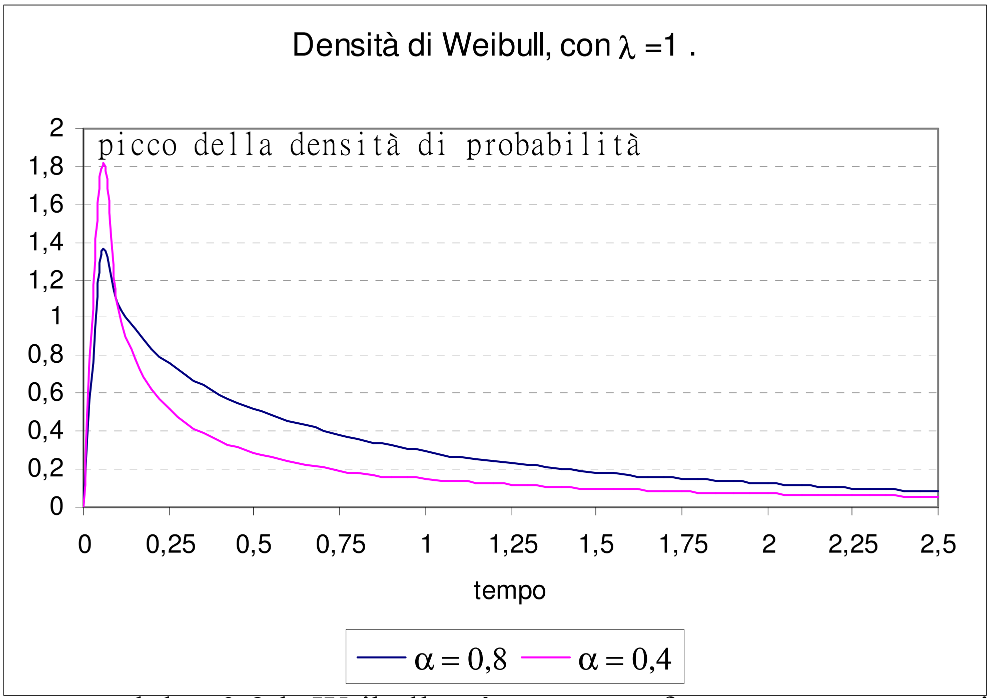
Porta alla definizione delle funzioni densità e distribuzione di Weibull :

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

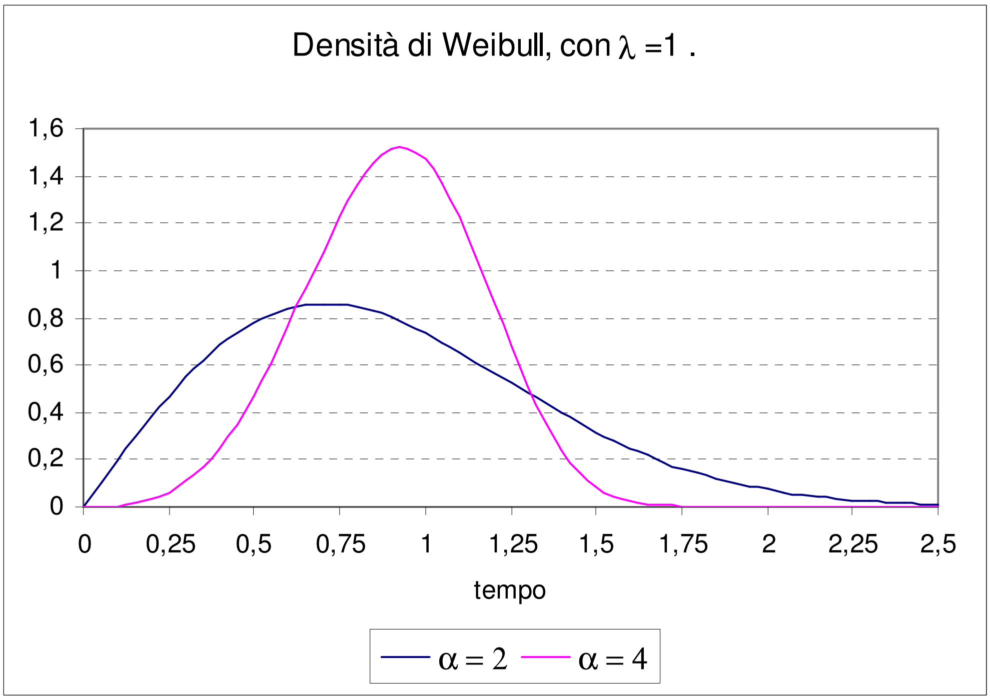
Weibull può descrivere la durata di vita per un fenomeno la cui "probabilità di morire" può variare nel tempo, a differenza della distribuzione esponenziale che prevede tassi di guasto costanti nel tempo.

Rappresentazione grafica del modello di Webull

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<=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,…,n,…] 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, ∞), 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[numero]=numero quindi
E[E[X]]=numero!

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

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

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

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

non si preoccupa dove
si discostano
gli scarti della differenza

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

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

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

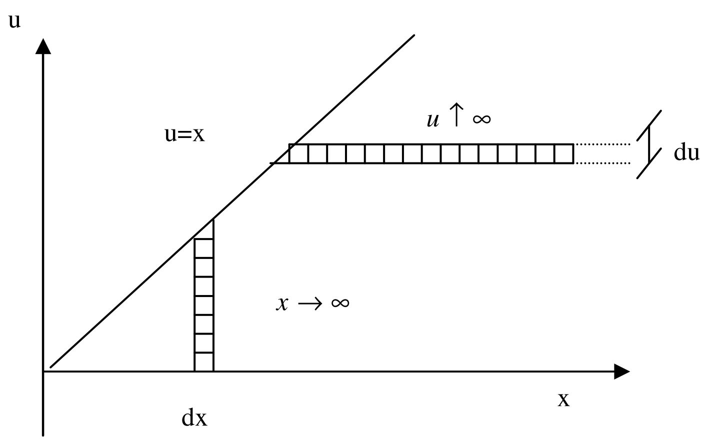
Per il calcolo del valore atteso di una variabile aleatoria non negativa, vale la seguente formula alternativa:

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

PROVA:

$$\int_{x=0}^{\infty} x f_X(x) dx = \int_{x=0}^{\infty} \int_{u=0}^{x} du f_X(x) dx =$$

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



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

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

Infatti:

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

Calcolo del valore atteso della legge esponenziale

Si possono seguire due metodi alternativi:

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

ovvero risolvendo per parti

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

Poiché:

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

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

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

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

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/λ. 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:

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

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

Dalla definizione:

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

ricordando che:

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

si ottiene:

$$E\left[X^2\right] = -\int_0^\infty t^2 R'(t)dt$$

$$= -t^2 R(t)\Big|_0^\infty + \int_0^\infty 2t^{2-1} R(t)dt = \int_0^\infty 2t\,R(t)dt$$

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

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

$$VAR[X] = E\left[X^2\right] - E[X]^2 \qquad \text{(Utile in futuro)}$$

(varianza è momento del secondo ordine meno momento del primo ordine al quadrato)

$$= \int_0^\infty 2t\,R(t)dt - \left(\int_0^\infty R(t)dt\right)^2$$

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

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 $\left(\beta \triangleq \frac{1}{d^{1/\alpha}} \right)$ secondo notazione di Excel.

Allora risultano:

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

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

(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:

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

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

Poniamo $z = \alpha x^\beta$

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

Da qui $\Rightarrow$ $$E[X] = \int_{z_1=0}^{\infty} \alpha\beta \frac{z_1}{\alpha} \cdot e^{-z_1} \cdot \frac{1}{\alpha\beta} \cdot \left(\frac{z_1}{\alpha}\right)^{\frac{1}{\beta}-1} \cdot dz_1$$

$$= \int_0^{\infty} \left(\frac{z_1}{\alpha}\right)^{\frac{1}{\beta}} \cdot e^{-z_1} \cdot dz_1$$

$$= \alpha^{-\frac{1}{\beta}} \int_{z_1=0}^{\infty} z_1^{\frac{1}{\beta}} e^{-z_1} \cdot dz_1$$

Ebbene, l'integrale indefinito che vedete è la funzione "gamma" di Eulero (precisamente: il II integrale di Eulero) che si riporta sui libri con la seguente notazione: $$\Gamma(\alpha) \triangleq \int_{z=0}^{\infty} z_1^{\alpha-1} \cdot e^{-z_1} \cdot dz_1$$
con $\alpha > 0$

Il calcolo di $\Gamma(\alpha)$ per specifici valori di α è detto "guaio di Eulero" ☺

Nel nostro caso, ci limitiamo ad osservare quanto segue:

$$\int_{z=0}^{\infty} z_1^{\frac{1}{\beta}} \cdot e^{-z_1} \cdot dz_1 \equiv \Gamma\left(1+\frac{1}{\beta}\right) \Rightarrow \text{conclusione:}$$

$$\Rightarrow E[X] = \alpha^{-\frac{1}{\beta}} \cdot \Gamma\left(1+\frac{1}{\beta}\right)$$

Statistiche dell'Ordinamento

Date $X_1, X_2, ..., X_n$ variabili aleatorie indipendenti ed identicamente distribuite, sono utili nell'analisi probabilistica le variabili aleatorie $Y_1,...,Y_n$ definite alla seguente maniera:

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

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

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

Y_i sono funzioni di variabili aleatorie e note come statistiche dell'ordinamento. Infatti, detta Y_k la statistica di ordine k *(k=1,...,n)*, è immediato riconoscere che, se X_i rappresenta il tempo al guasto del componente *i (i=1,...n)*, allora la statistica di ordine $k = n\text{-}m\text{+}1$ 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 $R_1(t),....,R_n(t)$ 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 risulti funzionante all'istante *t* un campione di *j* componenti indipendenti ; In particolare probabilità congiunta 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:

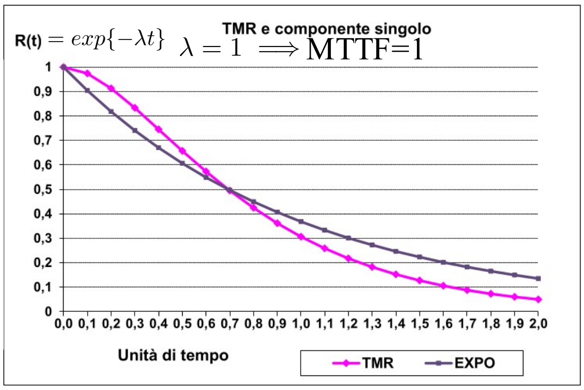
$$R_{m/n}(t) = \sum_{j=m}^{n} \binom{n}{j} [1 - R(t)]^{n-j} R(t)^j$$

Casi particolari

- **Il TMR con componenti identici** ($R_1(t) = ... = R_n(t) \triangleq R(t)$). "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":

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

$$R_{1/n}(t) = \sum_{j=1}^{n}\binom{n}{j}[1-R(t)]^{n-j}R(t)^j = \sum_{j=0}^{n}\binom{n}{j}[1-R(t)]^{n-j}R(t)^j - \binom{n}{0}[1-R(t)]^n = [1-R(t)+R(t)]^n - 1\cdot[1-R(t)]^n$$

$= 1-[1-R(t)]^n$ esplicitando le mortalità: $[1-R(t)]^n \implies = 1 - Pr\{X_1 \le t \cap \dots \cap X_n \le t\} = 1 - \prod_{i=1}^{n} F_{X_i}(t)$ Eguagliando i due

risultati arancioni $F_{Y_n}(t) = \prod_{i=1}^{n} F_{X_i}(t)$ perciò la funzione tempo al guasto del ultimo è dato dal prodotto delle distribuzioni.

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

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

Eguagliando i risultati in fucsia, ed esplicitando R(t) si ottiene che: $\prod_{i=1}^{n}[1-F_{X_i}(t)] = 1 - F_{Y_1}(t)$

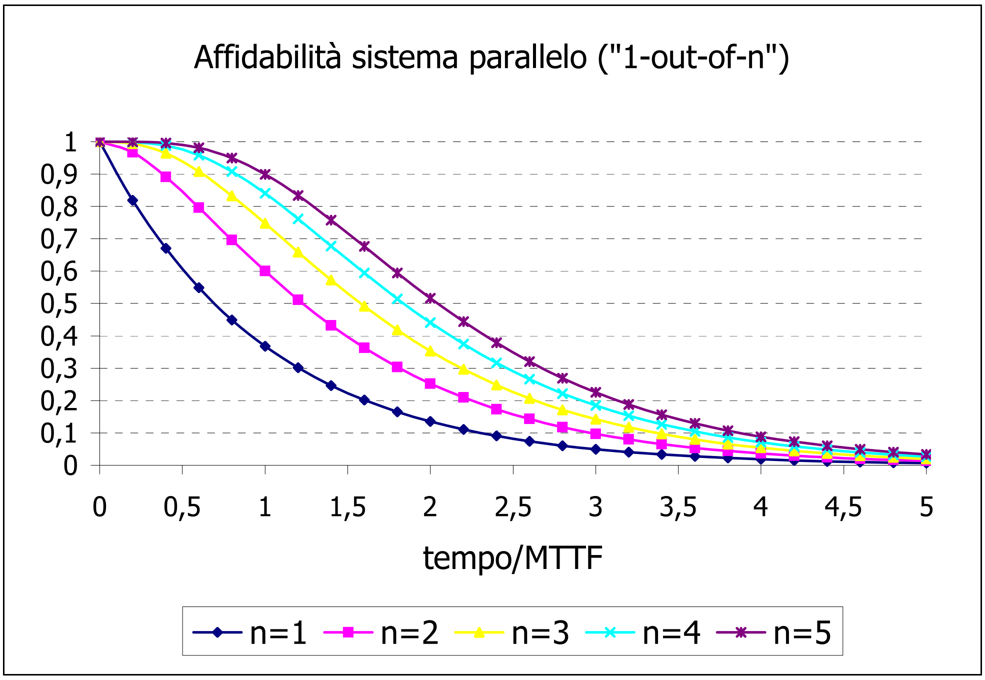
ovvero che la produttoria fra le complementari è pari alla sopravvivenza.

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:

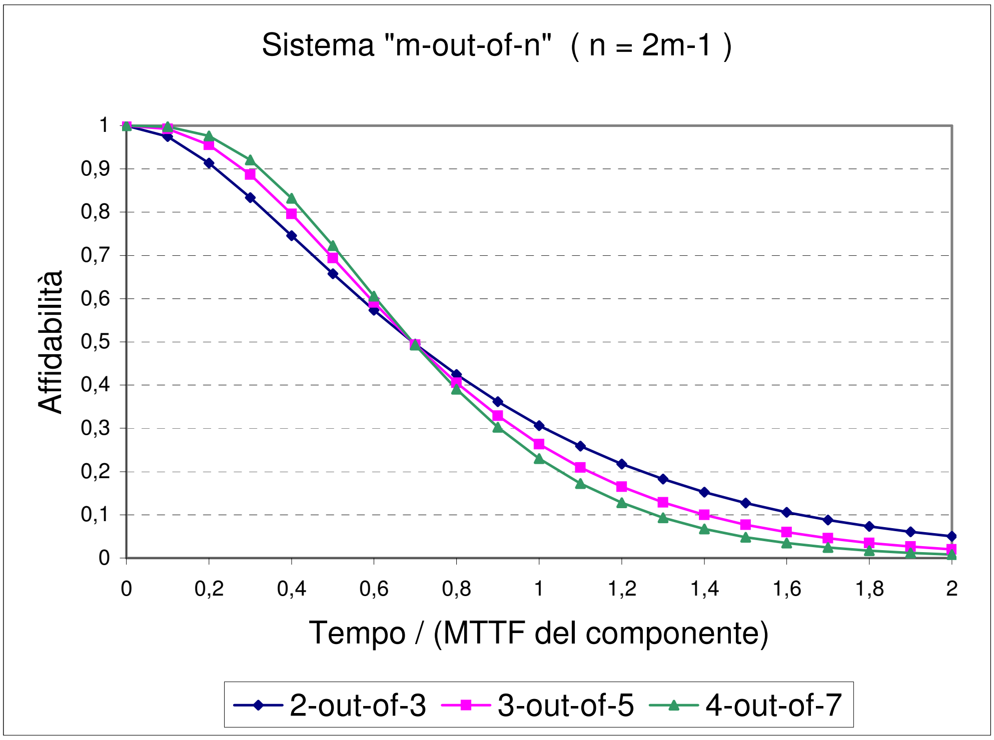
$$F_{X_i}(t) = 1 - e^{-\lambda_i t} \implies F_{Y_1}(t) = 1 - \prod_{i=1}^{n} e^{-\lambda_i t} = 1 - exp[-\sum_{i=1}^{n}\lambda_i]t$$

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

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

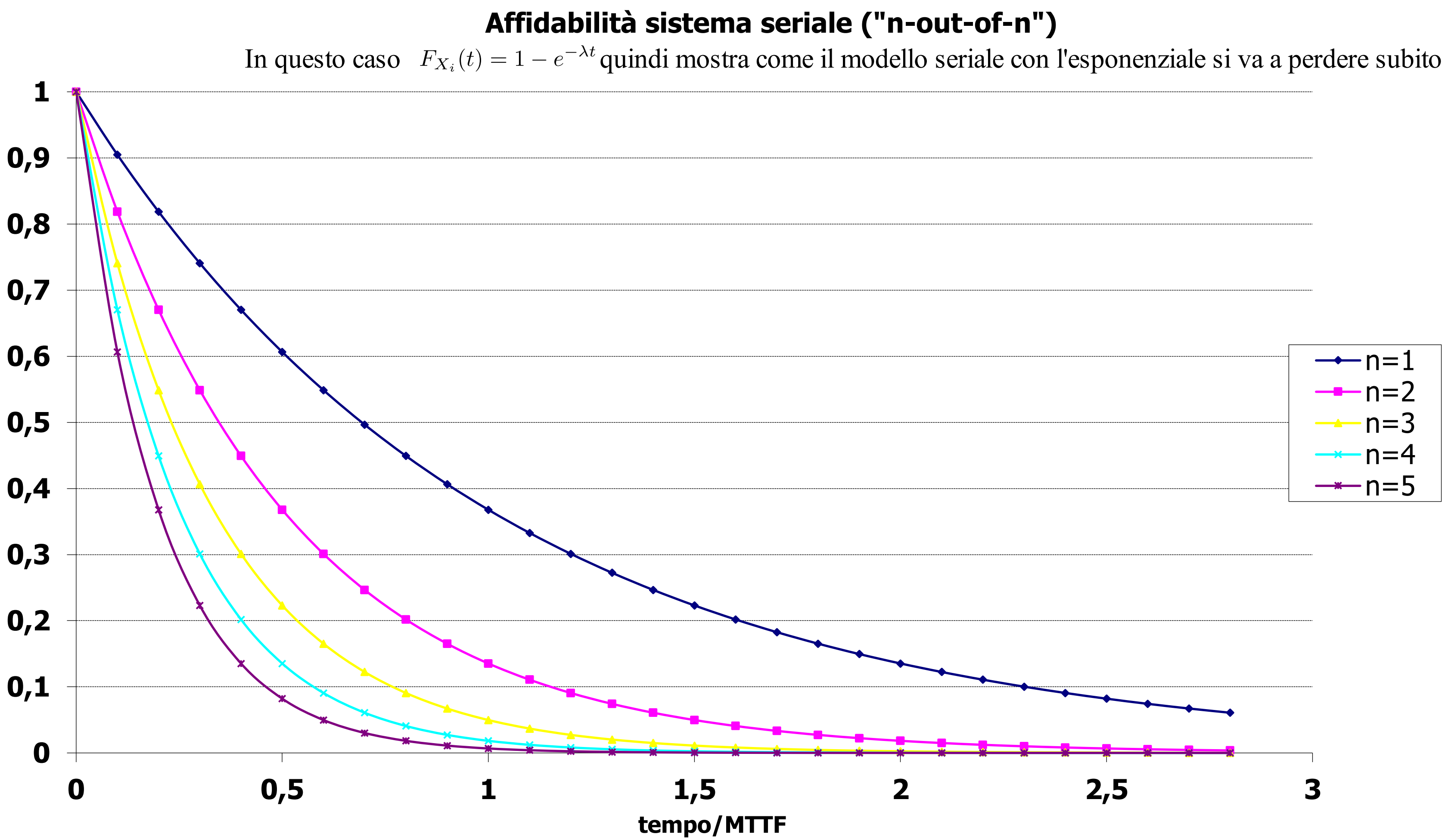


dopo 4 componenti in parallelo l'affidabilità che si guadagna non è poi cosi tanta



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



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:

$$\lambda_i = {}^{1}\!/_{10000} = 10^{-4}\, ore^{-1} \quad \forall i = 1,\ldots,4$$

ossia, $\lambda_i = \lambda$ *(i=1, …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 $\lambda_i = \lambda$ $(i=1, \ldots 4)$ 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:

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

$$=4\cdot\left[R(t=500)\right]^{3}-3\cdot\left[R(t=500)\right]^{4}$$

$$=0.986$$

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:

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

$$=1\cdot 1\cdot\left[1-e^{-10^{-4}\cdot 500}\right]^{2}+2\cdot\left[e^{-10^{-4}\cdot 500}\right]\cdot\left[1-e^{-10^{-4}\cdot 500}\right]+$$

$$+1\cdot\left[e^{-10^{-4}\cdot 500}\right]\cdot 1\quad=0.013$$

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

$\Rightarrow$ 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'esponenzile è costante

Il *"rischio di guasto"* di un componente sta crescendo linearmente col tempo d'uso, secondo un coefficiente pari a 0,2 (u.t.$^{-2}$). 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

$$R(u)\Big|_{u=2} = P(X \geq 2) = \exp\left[-\int_{t=0}^{2} h(t)dt\right] =$$

$$= \exp\left[-\int_{t=0}^{2} 0.2t dt\right] = \exp(-0.1 \cdot u^2)\Big|_{u=2} = 0.6703$$

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:

Per la probabilità uso f(t)dt
per il tasso uso h(t)dt

F è differenziabile per ipotesi

$$P(X=t)=dF(t)=f(t)dt \quad \overset{\text{se}}{con}\ dt=0 \quad \Rightarrow \quad P(X=t)=0$$

La probabilità di un guasto in [t; t+dt] si calcola come:

$$P(t \le X \le t+dt) = f(t)\cdot dt$$

ricordando la relazione tra f(t) e R(t)

dove

$$f(t) = -R'(t) = h(t)\cdot R(t) = \quad 0.2t\cdot e^{-0.1t^2}$$

R(t) è la condizionante
h(t)dt è la condizionata

$$con \quad dt = 0.01 \Rightarrow$$

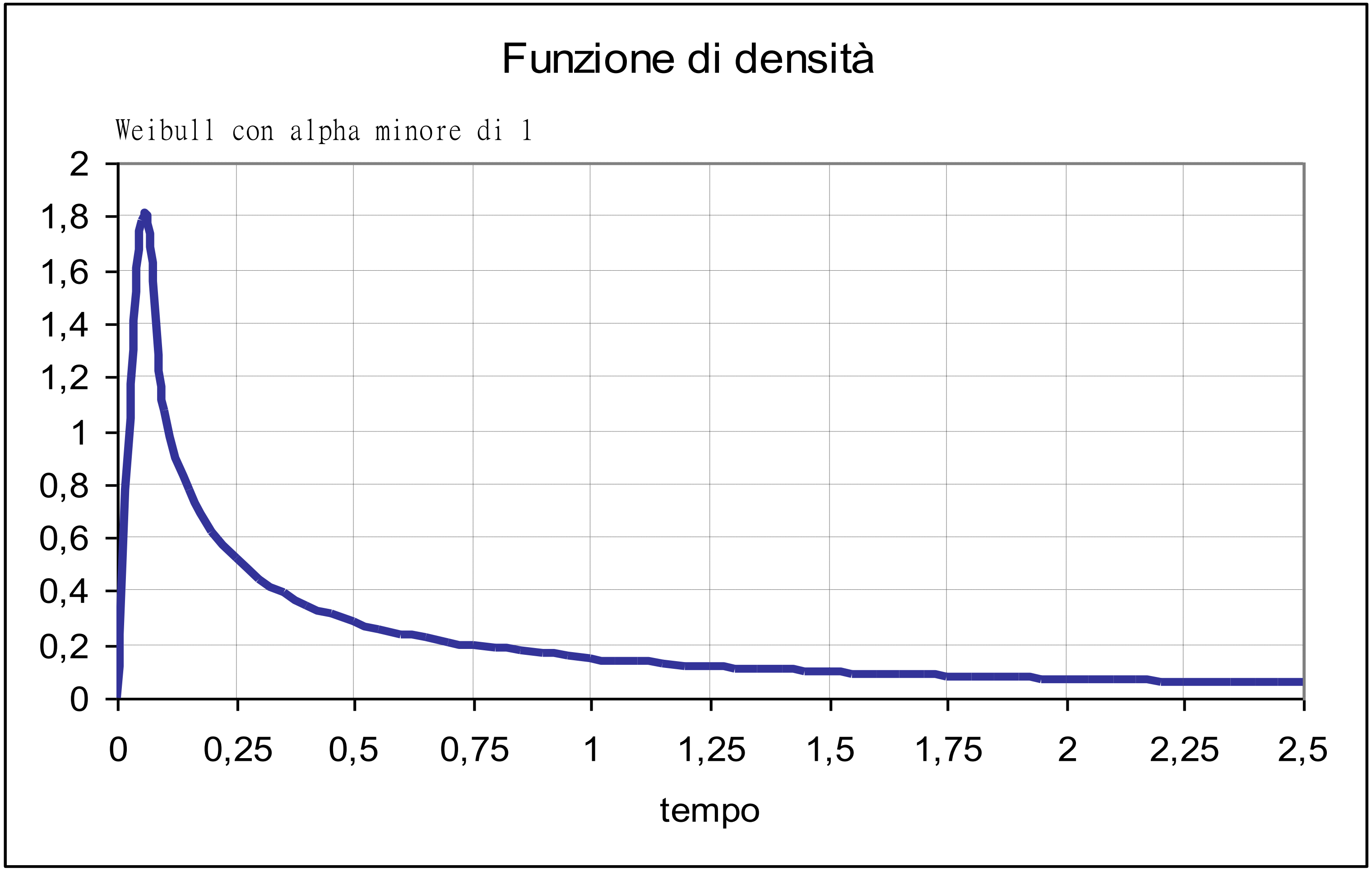
$f(t)dt = h(t)dt$ se R(t)=1 ovvero è certa

$$f(t)dt = f(0.25)\cdot 0.01 = 0.000497$$

Quantità trascurabile

DIPENDE
DAL CONTESTO

Esercizio 1
(Considerazione sul quesito 3…)



Esercizio 2
(Descrizione del problema)

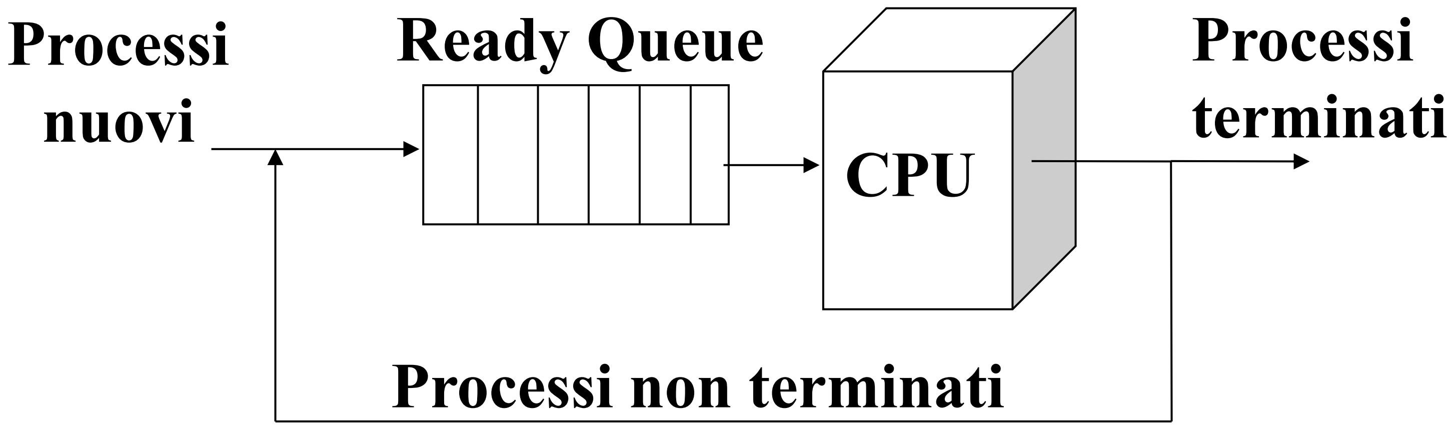
Prova a riformularlo in un altro contesto

- 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 <u>più di un solo</u> *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)

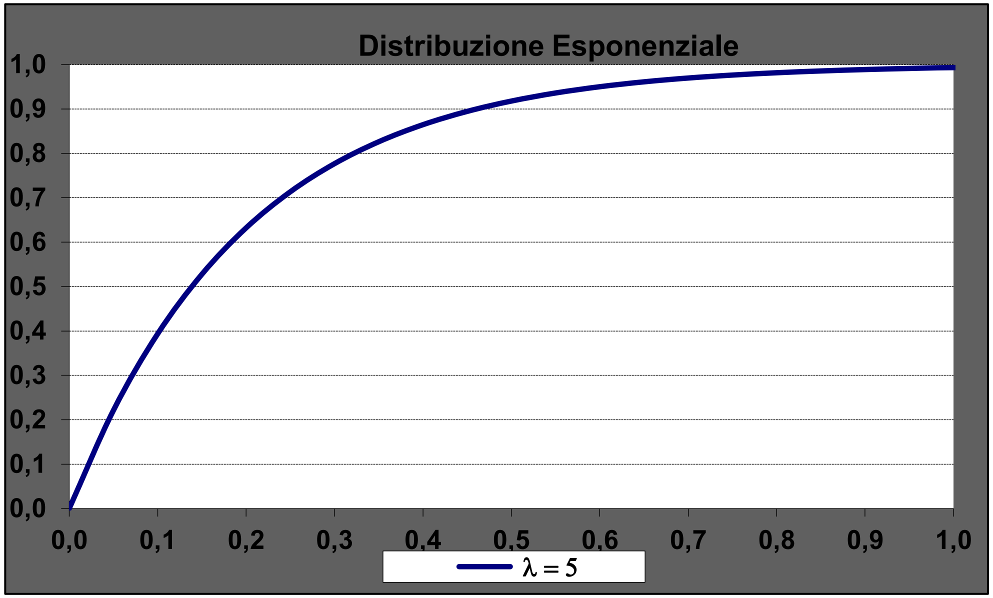
- Avendo solo la media (valore atteso) del tempo di esecuzione, il modello da adottare è quello esponenziale.

Sia X “la durata del tempo di esecuzione”

Allora: $$E[X] = \frac{1}{\lambda} = \frac{2}{10} \Rightarrow \lambda = 5\sec^{-1}$$

e quindi

$$F_X(t) = 1 - e^{-\lambda t}$$
$$= 1 - e^{-5t}$$



Esercizio 2
(Soluzione dei quesiti 2-4)

- La probabilità che sia sufficiente un unico *time slice* è:

$$P\left(X \le 0.02\right) = \int_{0.0}^{0.2} f_X(u)du = F_X(0.02) = 0.095$$

- La probabilità che ne occorra più di uno solo è:

$$P(X > 0.02) = 1 - P(X \le 0.02) = 0.905$$

- La probabilità che occorrano da 1 a 3 *time slice* è:

non uso il differenziale perché non è un tempo piccolo l'intervallo

$$P\left(0.02 \le X < 0.06\right) = F_X(0.06) - F_X(0.02) = 0.164$$

- La probabilità che venga completato col suo secondo *time slice*, posto che non ce l'abbia fatta col suo primo.

$$P\left(X \le 0.04 \mid X > 0.02\right) = P\left(X \le 0.02\right) = 0.095$$

(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

$$= \frac{P(\bar{x} \leq X \leq \bar{x} + y)}{P(X \geq \bar{x})} = \frac{F(\bar{x} + y) - F(\bar{x})}{1 - F(\bar{x})}$$

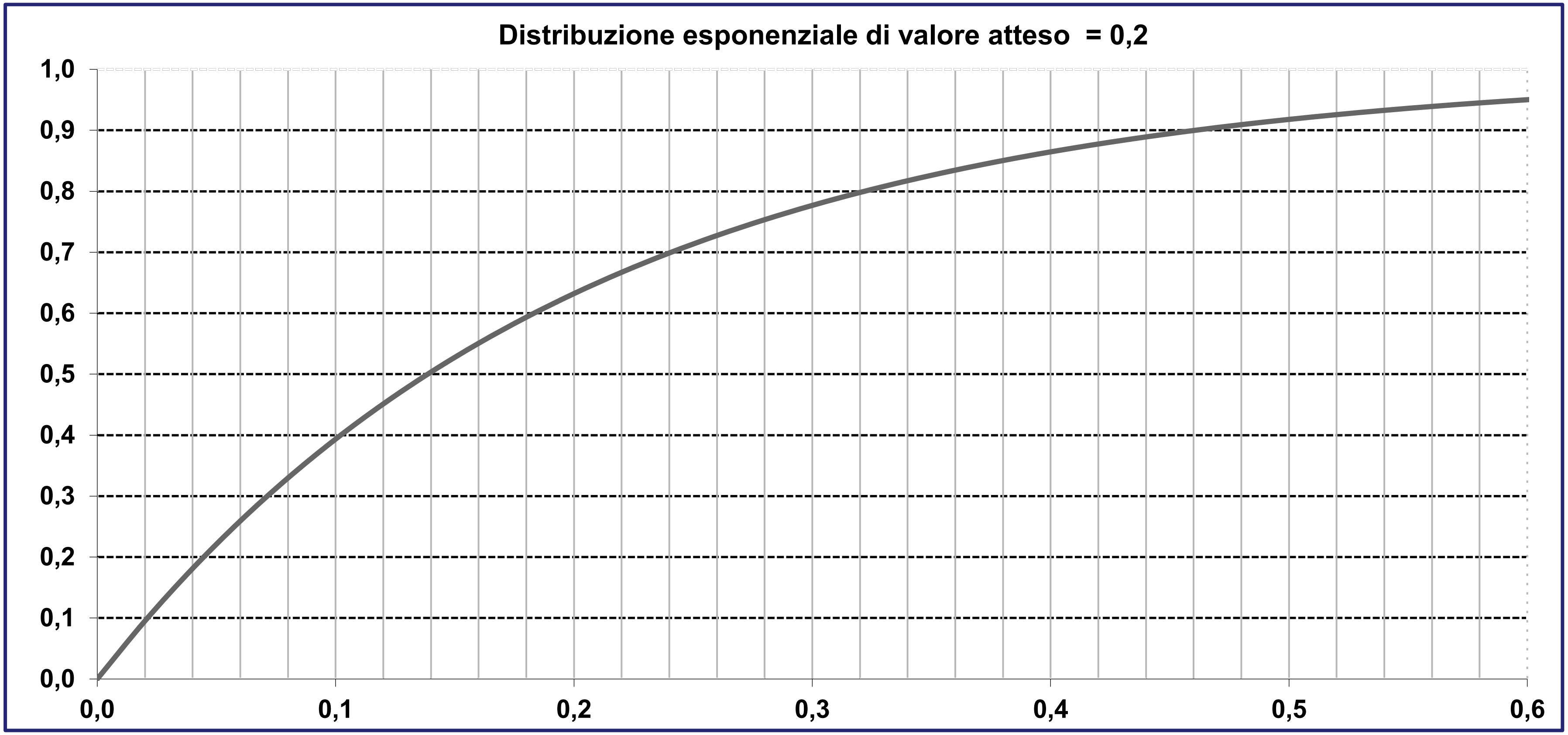
$$= \frac{F(0.02 + 0.02) - F(0.02)}{1 - F(0.02)} = 0.095$$

DOMANDA extra:

A cosa corrisponde la seguente?

$$\boxed{\tilde{F}(x) \hat{=} \frac{F(x)}{F(\bar{x})}, \; 0 \leq x \leq \bar{x}.}$$

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:

$$X \hat{=} max\{X_1,\ldots,X_n\}$$

  - Allora:

$$F_X(t) = Pr\left(X_1 \le t,\ldots,X_n \le t\right) = \prod_{i=1}^{n} F_{X_i}(t)$$

  - L'affidabilità aumenta assai poco con *n*:

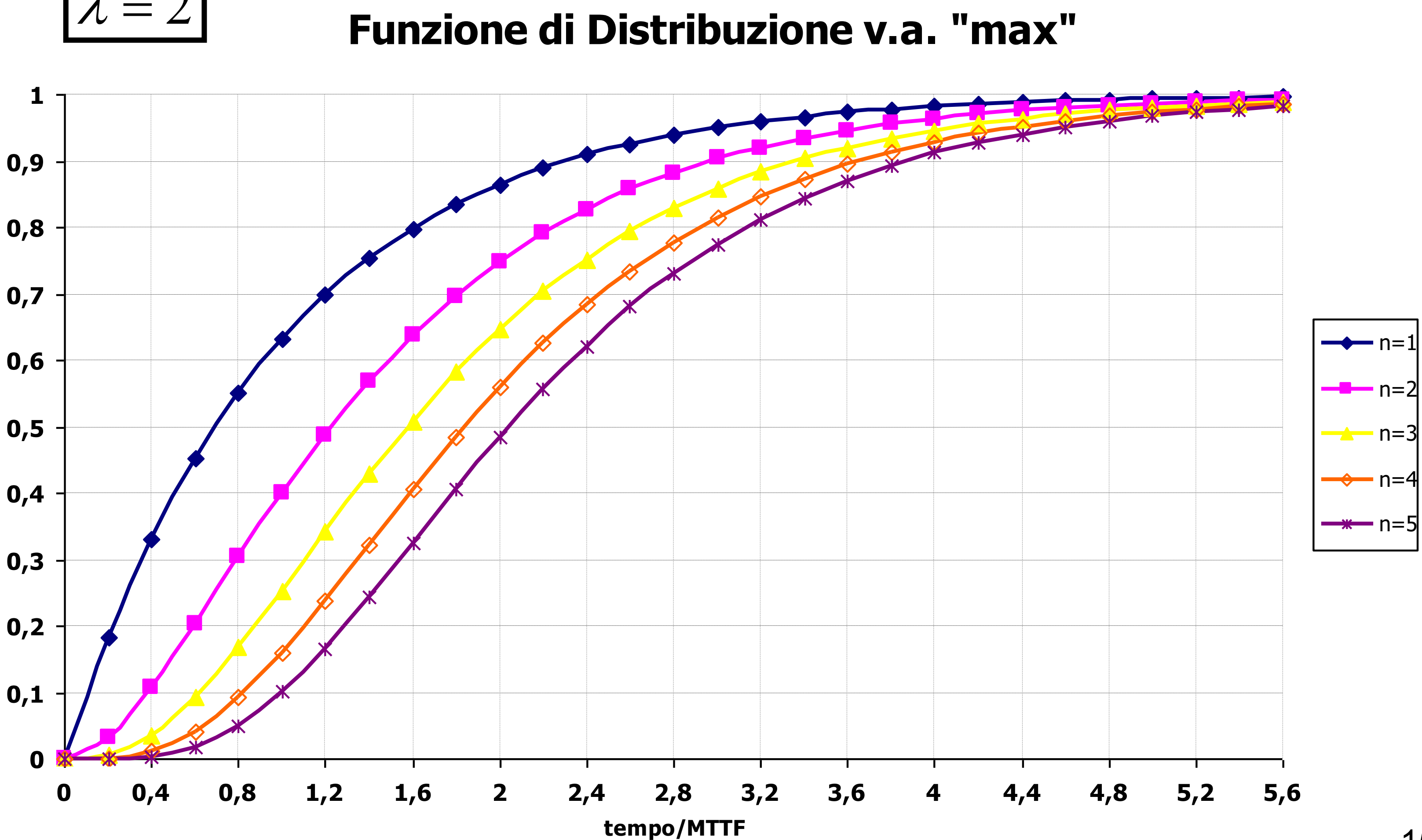
$$\boxed{F_X(t) = \left(1 - e^{-\lambda t}\right)^n}$$

Caso "esponenziale"

RICAP.: distribuzione della v.a. “max” (2_di_2)

Curve della $F_X(t)$ al variare di n

$\boxed{\lambda = 2}$



RICAP.: distribuzione della v.a. “min” (1_di_2)

- La v.a. “min” è:

Allora: $$X \hat{=} min\left\{X_1, \ldots, X_n\right\}$$

$$1 - F_X(t) = \Pr\left(X_1 > t, \ldots, X_n > t\right) = \prod_{i=1}^{n} 1 - F_{X_i}(t)$$

L’affidabilità diminuisce assai velocemente con *n*:

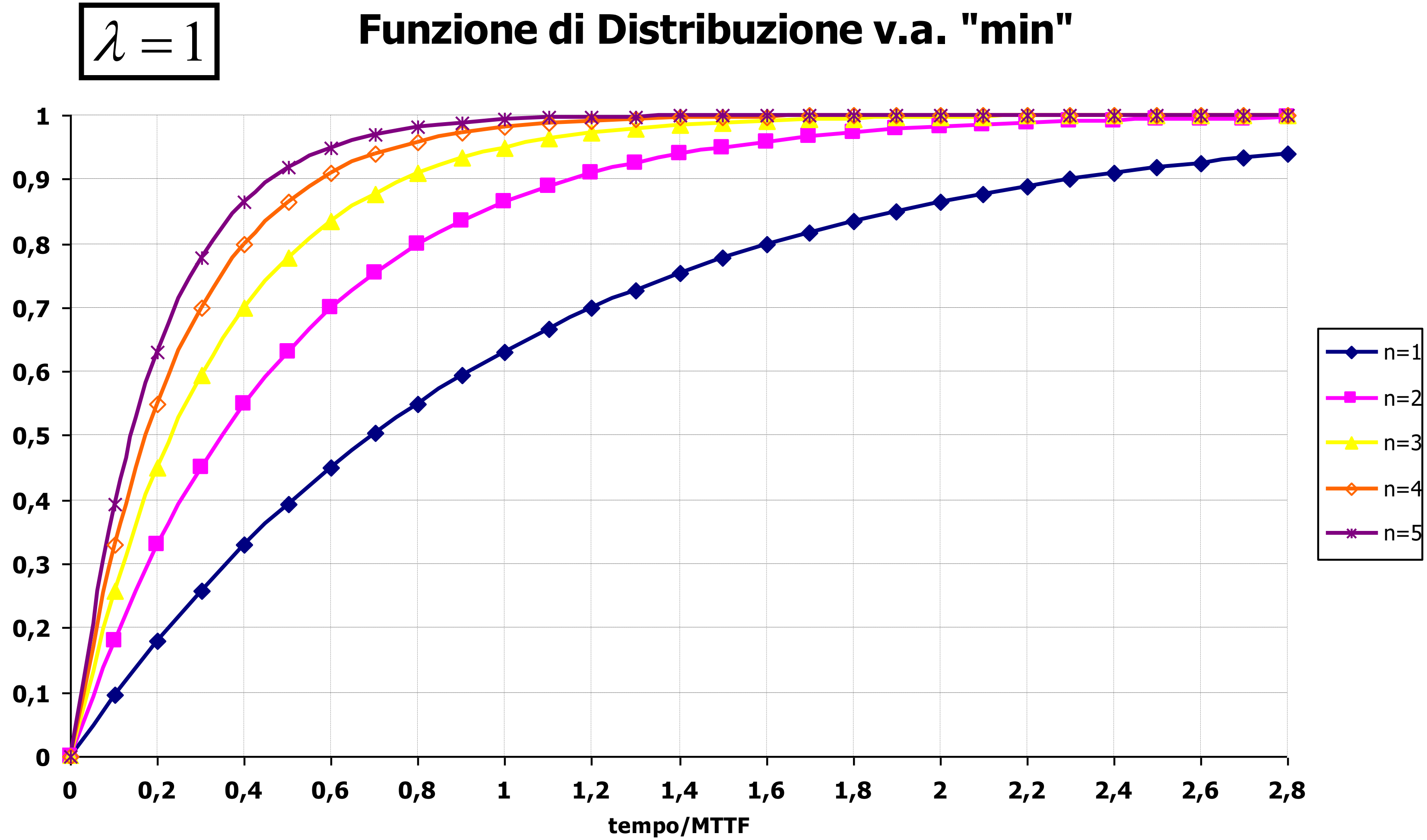
$$\boxed{F_X(t) = 1 - (e^{-\lambda t})^n}$$

Caso “esponenziale”

RICAP.: distribuzione della v.a. “min” (2_di_2)

Curve della $F_X(t)$ al variare di n

$\boxed{\lambda = 1}$



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:

$$X \hat{=} max\{X_1, X_2\}$$

  - 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)

$$F_X(t) = \left(1 - e^{-\mu_1 t}\right)\left(1 - e^{-\mu_2 t}\right)$$

$$= 1 - \left(e^{-\mu_1 t} + e^{-\mu_2 t} - e^{-(\mu_1 + \mu_2)t}\right)$$

Esercizio 4
(Soluzione al quesito 2)

CALCOLO

- Il valore atteso della durata del blocco:

$$E[X] = \int_{t=0}^{\infty} \left(1 - F_X(t)\right) dt$$ che tra l'altro è l'MTTF. Appunto integrale della sopravvivenza

$$= \int_{t=0}^{\infty} \left(e^{-\mu_1 t} + e^{-\mu_2 t} - e^{-\left(\mu_1 t + \mu_2 t\right)}\right) dt$$

$$= \frac{1}{\mu_1} + \frac{1}{\mu_2} - \frac{1}{\mu_1 + \mu_2} \cong 7.5\, sec$$

cos'é 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:

$$X \hat{=} min\{X_1, X_2\}$$

  - 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 :

$$1 - F_X(t) = e^{-\mu_1 t} \cdot e^{-\mu_2 t}$$

valore atteso

$$F_X(t) = 1 - e^{-(\mu_1 + \mu_2)t} \quad ; \quad E[X] = \frac{1}{\mu_1 + \mu_2} = 1.33\, sec$$

Richiamo formule sistemi “m-out-of-n”

- Durate indipendenti e identiche, con stessa legge esponenziale e durata media = 1/λ :

$$R_{1/n}^{\text{parallelo}}(t)=\sum_{j=1}^{n}\binom{n}{j}\left[e^{-\lambda t}\right]^{j}\left[1-e^{-\lambda t}\right]^{n-j}=1-\left[1-e^{-\lambda t}\right]^{n}$$

$$R_{m/n}(t)=\sum_{j=m}^{n}\binom{n}{j}\left[e^{-\lambda t}\right]^{j}\left[1-e^{-\lambda t}\right]^{n-j},\quad 1<m<n$$

$$R_{n/n}^{\text{seriale}}(t)=\sum_{j=n}^{n}\binom{n}{j}\left[e^{-\lambda t}\right]^{j}\left[1-e^{-\lambda t}\right]^{n-j}=\left[e^{-\lambda t}\right]^{n}$$

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).

La porzione di testo in blu è una ipotesi semplificativa
che varrà sempre ed è molto semplificativa

che modello si può usare? quale no?

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: $E[X] = \frac{1}{\lambda} = 60min \implies \lambda = \frac{1}{E[X]} = \frac{1}{60\ min} = 0.017min^{-1} = 0.00027778\ s^{-1}$

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 tavolo 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.

Inizio analisi statistica

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 fondato 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 $E[X]=$ 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 distrosione". Nella standard dopo 3 deviazione, 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

per z che tende a infinito
la funzione tende a 0

Con le ulteriori condizioni:

funzione crescente per z<0
funzione decrescente per >0

$$\frac{d\, f_Z(z)}{dz} = \begin{cases} >0 & z<0 \\ <0 & z>0 \end{cases} \qquad \text{e} \qquad f_Z(z \to \pm\infty) \to 0$$

Più:

concava

condizione per
punto di massimo o minimo

$$\left.\frac{d\, f_Z(z)}{dz}\right|_{z=0} = 0 \qquad \text{e} \qquad \left.\frac{d^2 f_Z(z)}{dz^2}\right|_{z=0} < 0$$

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 → $\frac{d\, f_Z(z)}{dz} = -f_Z(z)\, z;$ con $f_Z(z) > 0$ per $-\infty < z < +\infty$

Integrando la precedente si ottiene:

$$f_Z(z) = k \exp\left\{-z^2/2\right\}, \quad \text{con} \quad k>0$$

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 e 0 +infinito è perché è simmetrica

Preso il secondo integrale di Eulero $\int_{t=0}^{+\infty} t^{\alpha-1} exp\{-t\} dt = \Gamma(\alpha)$

$$1 = \int_{-\infty}^{+\infty} k \exp\left\{-z^2/2\right\} dz = k \frac{2}{\sqrt{2}} \int_{t=0}^{+\infty} t^{\frac{1}{2}-1} \exp\{-t\} dt \qquad \left(t \triangleq z^2/2\right)$$

si ottiene il risultato

e poiché riconosciamo la "funzione gamma" nell'ultimo integrale indefinito:

$$\int_{t=0}^{+\infty} t^{\frac{1}{2}-1} \exp\{-t\} dt = \Gamma(1/2) = \sqrt{\pi}$$

concludiamo che la densità normale standard è:

rispetto alla non standard al denominatore
sigma non c'è, perché pari a 1

$$f_Z(z) = \frac{1}{\sqrt{2\pi}} \exp\left\{-z^2/2\right\}, \qquad -\infty < z < +\infty$$

e da qui la ***distribuzione normale standard***, in forma di funzione integrale:

$$F_Z(z) = \int_{-\infty}^{u=z} \frac{1}{\sqrt{2\pi}} \exp\left\{-u^2/2\right\} du, \qquad -\infty < z < +\infty \text{ (rimane funzione integrale)}$$

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 ed un dato valore u > 0, i.e., $P(-\infty \le z \le u)$

Osservazione: $F_Z(-z) = 1 - F_Z(z)\ z \to 0$ 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

$$\begin{aligned} I^2 &= \int_{-\infty}^{\infty} \frac{1}{2\pi} e^{-(1/2)x^2}\, dx \cdot \int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi}} e^{-(1/2)y^2}\, dy \\ &= \frac{1}{2\pi} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} e^{-(1/2)(x^2+y^2)}\, dx\, dy \end{aligned} \tag{6.4.6}$$

Substituting $x = r\cos(\theta)$ and $y = r\sin(\theta)$ in Eq. (6.4.5) we obtain,

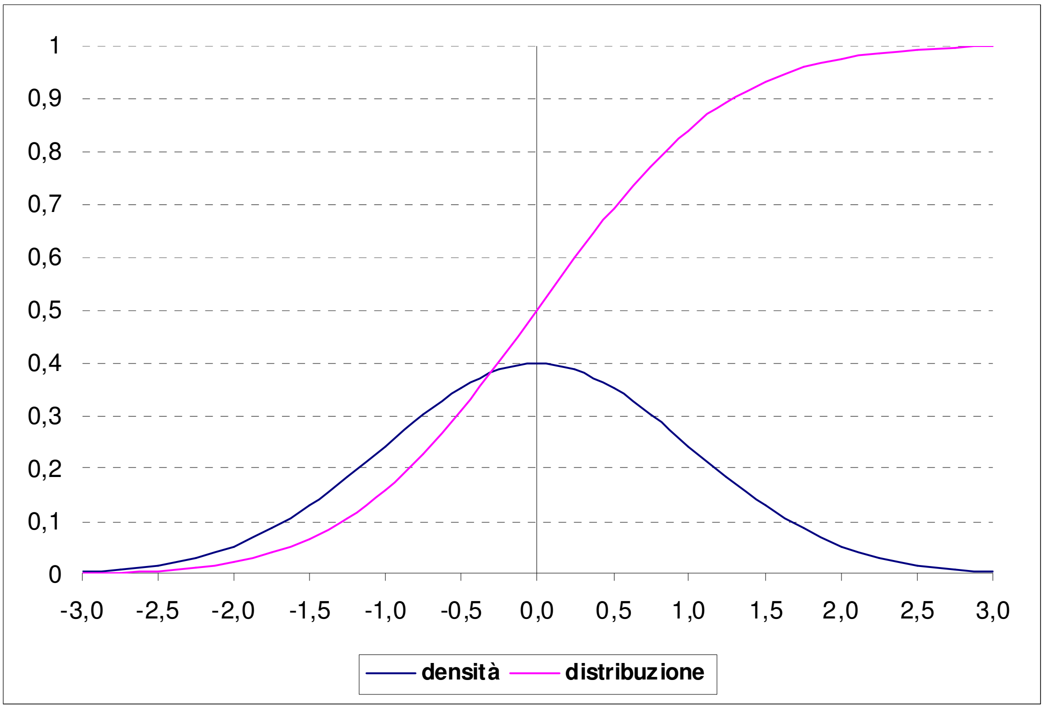
$$I^2 = \frac{1}{2\pi} \int_0^{2\pi} \int_0^{\infty} r e^{-(1/2)r^2}\, dr\, d\theta = 1 \tag{6.4.7}$$

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

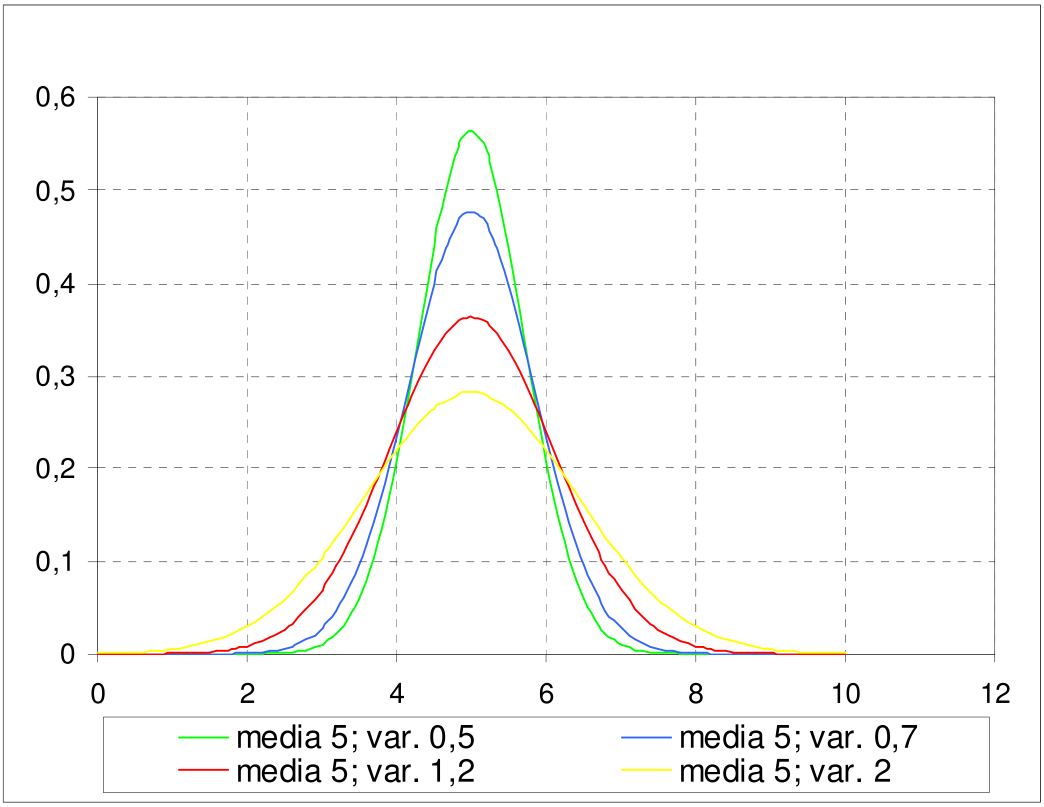
Rappresentazioni della normale

è "larga"
3 deviazioni standard poi tende a 0



in blu
la famosa
campana di Gauss

La normale standard.



Esempi di normale NON standard.

DETTAGLIO TECNICO_1

dimostrazione del perché esce sigma al quadrato nella varianza

PER DEFINIZIONE DI VARIANZA:

$$VAR[X] \triangleq E\left[(X - E[X])^2\right]$$

ALLORA:

$$VAR[aX] = E\left[(aX - E[aX])^2\right]$$

$$= E\left[(aX - aE[X])^2\right]$$

$$= E\left[a^2 \cdot (X - E[X])^2\right]$$

$$= a^2 E\left[(X - E[X])^2\right]$$

$$= a^2 \, VAR[X]$$

La distribuzione “normale” (9/10)

$$Sia \quad X \hat{=} \sigma Z + \mu, \quad con \quad -\infty < \mu < +\infty \quad e \quad \sigma > 0$$

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

$$VAR[X] = \sigma^2 VAR[Z] + \underbrace{VAR[\mu]}_{=0} = \sigma^2$$

perchè costante

quelli de grafico precedente

I due punti di flesso corrispondono alle ascisse $\mu \pm \sigma$

creo una relazione tra la gaussiana standard ed una non standard

$$\Pr\{X \leq x\} = \Pr\left\{Z \leq \frac{x-\mu}{\sigma}\right\} \Rightarrow z = \frac{x-\mu}{\sigma}$$

questo verrà usato per la non standard

densità non standard sarà:

$$f_X(x) = \frac{1}{\sigma\sqrt{2\pi}} \exp\left\{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2\right\}, \qquad -\infty < x < +\infty$$

DETTAGLIO TECNICO_2_e_3

utile per la relazione tra standard
e non standard

$$\Pr\{X \le x\} = \Pr\left\{Z \le \frac{x-\mu}{\sigma}\right\} \quad \Rightarrow \quad z = \frac{x-\mu}{\sigma}$$

$$\boxed{\begin{aligned} F_X(x) &\triangleq \Pr\{X \le x\} = \Pr\{\sigma Z + \mu \le x\} = \\ &= \Pr\left\{Z \le \frac{x-\mu}{\sigma}\right\} \triangleq F_Z\left(\frac{x-\mu}{\sigma}\right) \end{aligned}}$$

INOLTRE:

si dimostra
nella lezione
di funzioni di
variabili aleatorie

$$\boxed{\begin{aligned} f_X(x) &\triangleq \frac{d}{dx} F_X(x) = \frac{d}{dx} F_Z\left(\frac{x-\mu}{\sigma}\right) \\ &= \frac{1}{\sigma} f_Z\left(\frac{x-\mu}{\sigma}\right) \end{aligned}}$$

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 $_i$ e varianza σ_i^2 ($i = 1,\dots,n$). Allora, introdotte, le costanti reali $a_1, a_2, \dots, a_n$, la variabile aleatoria

$$S_n \triangleq \sum_{i=1}^{n} a_i X_i$$ combinazione lineare

è distribuita con una legge normale di media

$$= \sum_{i=1}^{n} a_i \ _i \quad \text{e varianza} \quad \sigma^2 = \sum_{i=1}^{n} a_i^2 \sigma_i^2 .$$ 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 tende alla Gaussiana standard ma la sommatoria delle variabili X_i no

$$\tilde{Z}_n \triangleq \frac{\sum_{i=1}^{n} X_i - \sum_{i=1}^{n} \ _i}{\sqrt{\sum_{i=1}^{n} \sigma_i^2}}$$

Si sta standardizzando ad una Gaussiana

che ha, per costruzione, media nulla e varianza unitaria converge in forma alla normale standard al crescere di n, cioè:

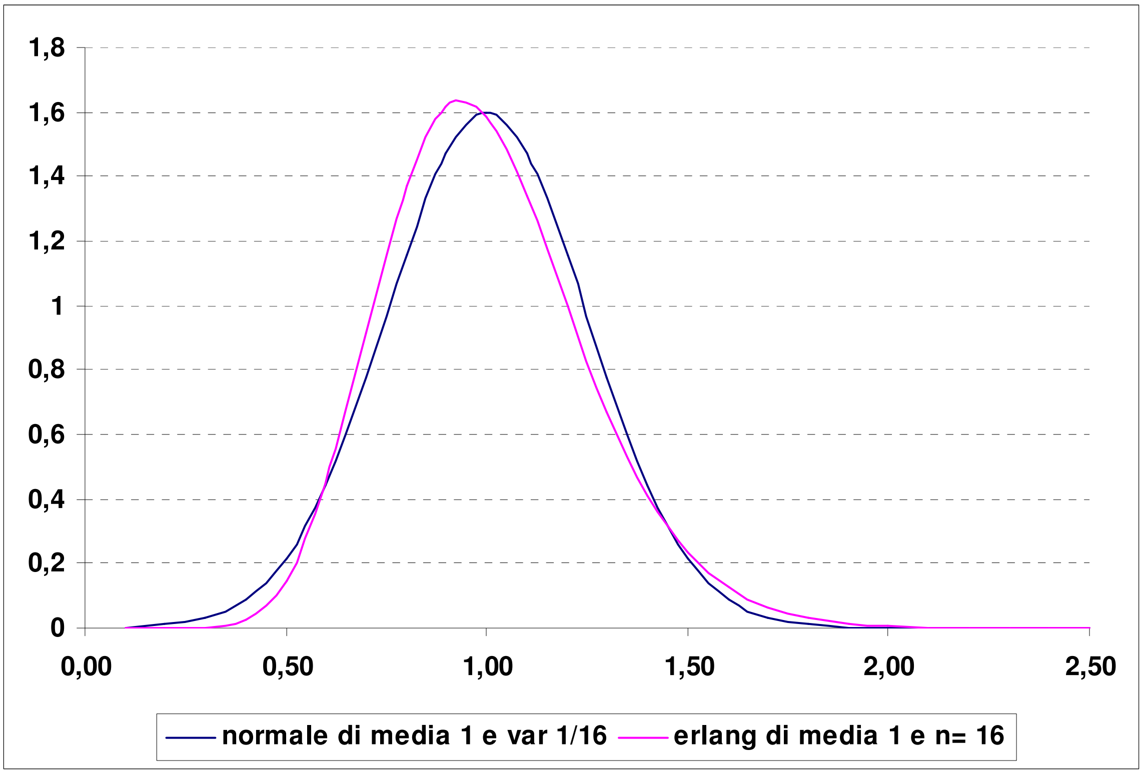
$$n \to \infty \Rightarrow \quad F_{Z_n}(z) \overset{\text{distr}}{\to} \int_{-\infty}^{u=z} \frac{1}{\sqrt{2\pi}} \exp\left\{-u^2/2\right\} du, \qquad -\infty < z < +\infty$$

già con n=30
si ha una normale

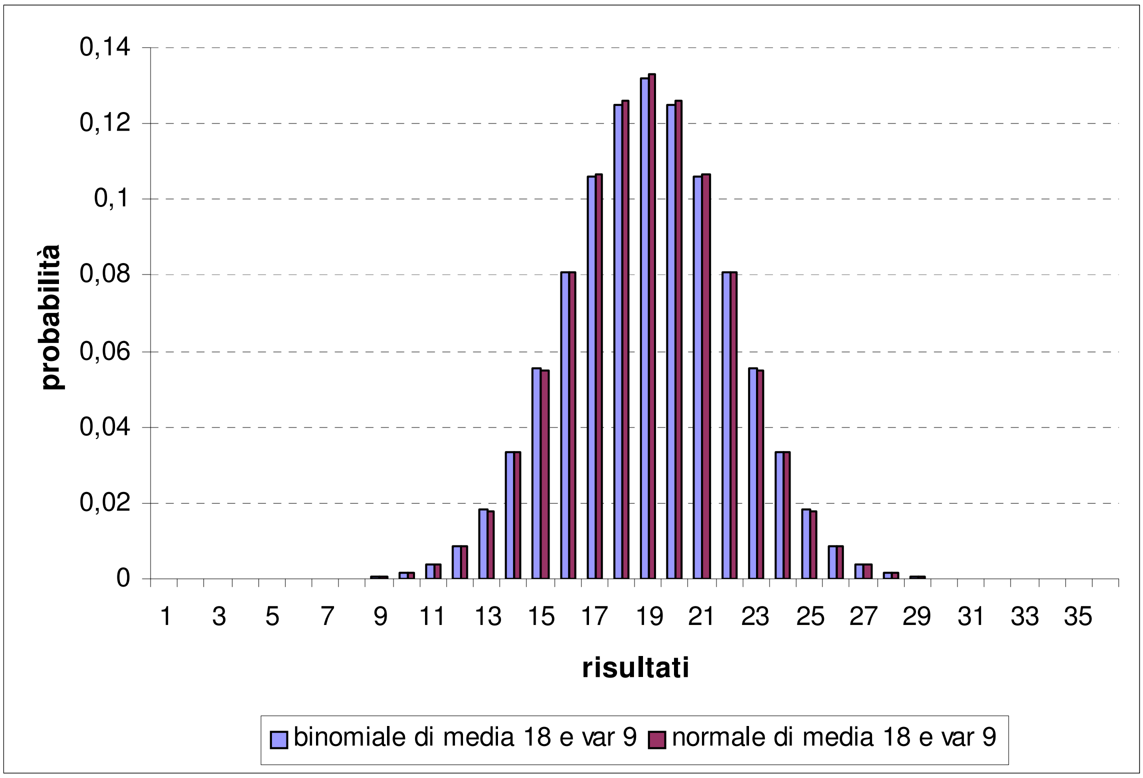
La dimostrazione dei due teoremi è tralasciata.

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

La funzione di *distribuzione congiunta*, di due var. al. (continue):

Caso d'uso:
tempi al guasto di due componenti soggetti sia a cause individuali di guasto sia a cause comuni di guasto

$$F_{X_1X_2}(t_1,t_2) \triangleq \Pr\{X_1 \le t_1 \cap X_2 \le t_2\}$$

$$0 \le t_1 < \infty,\ \ 0 \le t_2 < \infty$$

se ad esempio ho 2 processi in parallelo che devono conclude un job, la semplificazione sta nel rendere t1=t2.

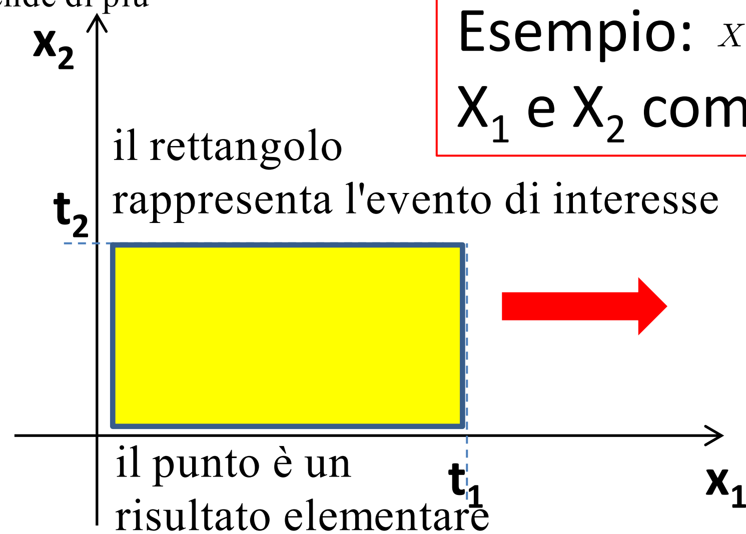
misura la probabilità congiunta che le durate (X_1 e X_2) delle due attività siano non superiori a t_1 e t_2, rispettivamente, quando non è lecito trattare X_1 e X_2 separatamente, l'una a prescindere dall'altra.

X_1, X_2 sono correlate (quale che sia)

Esempio: l'attesa in modello client-server è collegata alla durata dei servizi. Se un servizio dura di più si attende di più

Esempio: X_1, X_2 processi che comunicano nell'esecuzione

X_1 e X_2 come durate del viaggio di due amici, in R^2_+



VIAGGIO come ESPERIMENTO ALEATORIO.
DURATE DEL VIAGGIO come RISULTATI ELEM., cioè punti di coordinate (x_1, x_2).
AREA come EVENTO (CUMUL.) di INTERESSE

l'evento di interesse passa da un intervallo/punto (1 dimensione) ad un'area (2 dimensioni)

Ipotizzando di voler analizzare il comportamento di due processi che ogni lunedì compiono un job partendo assieme terminando però in tempi diversi. L'esperimento aleatorio è il compimento del job ogni lunedi. Il tempo di conclusione è la durata dell'attività che è misurata nel continuo. Ogni lunedi si hanno realizzazioni diverse delle variabili aleatorie. Come visto, l'area rappresentata sull'asse cartesiano è l'insieme di eventi elementari, in particolare ogni punto delle ascisse indica il tempo di conclusione del primo processo mentre ogni punto delle ordinate indica il tempo di conclusione del secondo processo.

$f(x_1, x_2)$ è la densità della congiunta che il primo processo si concluda tra x_1 e $x_1 + dx_1$ e che il secondo processo termini tra x_2 e $x_2 + dx_2$

$f(x_1, x_2)dx_1dx_2$ è invece la probabilità congiunta che due eventi accadano

La probabilità dell'evento di interesse come sappiamo, si calcola andando a sommare le singole probabilità degli eventi elementari che lo compongono. Essendo nel continuto ed essendo un'area la somma sarà espresse da un integrale doppio.

Cosa devo fare per escludere che uno o entrambi gli amici non rientrino prima di un certo tempo sicuramente?

Distribuzione marginale

La distribuzione congiunta può essere ottenuta a partire dalla definizione preliminare di una funzione densità congiunta, *f*,

$$F_{X_1X_2}(t_1,t_2)=\int_{u_1=0}^{t_1}\int_{u_2=0}^{t_2} f_{X_1X_2}(u_1,u_2)\,du_2\,du_1$$

In due dimensioni il volume (risultato dell'integrale) sarà unitario, differentemente dal caso monodimensionale dove è l'area ad essere unitaria.

Funzione di *distribuzione marginale*:

$$F_{X_i}^{(m)}(t_i)\hat{=}P(X_i\le t_i)\qquad i=1,2$$

probabilità che X_i finisca prima di t_i quale che sia lo stato di X_j

la m sta per marginale che si ricava e non è un modello primitivo

A PRESCINDERE DA X_J !

Quale che sia

Supponendo di conoscere la distribuzione congiunta delle durate dei rispettivi viaggi dei due amici, a cosa può servire la distribuzione marginale?

a calcolare la probabilità che un amico finisca il viaggio entro t_i, quale che sia la durata del viaggio del altro

Distribuzione congiunta e distribuzione marginale

Qualora sia data solo la densità congiunta, la distribuzione marginale della X1 si ottiene così:

Per ottenere la marginale di X1, occorre portare all'infinito X2 in questo modo integro su tutto lo spazio dei valori di X2 e resta tutto lo spazio di X1 che è indipendente da X2

$$\lim_{t_2 \to \infty} F_{X_1 X_2}(t_1, t_2) = \int_{u_1=0}^{t_1} \int_{u_2=0}^{\infty} f_{X_1 X_2}(u_1, u_2) du_2 du_1 = F_{X_1}(t_1)$$

Si conviene di definire densità marginale della X1 la seguente:

che integrata tra $u_1 = 0, t_1$ ritorna la distribuzione marginale F_{X_1}

$$F_{X_1(t_1)} \triangleq \int_{u_1=0}^{t_1} f_{X_1}(u_1) du_1$$

$$\boxed{f_{X_1}(u_1) \hat{=} \int_{u_2=0}^{\infty} f_{X_1 X_2}(u_1, u_2) du_2 \quad < \infty}$$

quindi la marginale è un caso limite dove una densità è definita come integrale indefinito!

Se converge ad un valore finito, si è associato un valore a u_1. Se ottengo più valori potrei generare una funzione.

Ripartizione congiunta e marginale nel discreto

RIFERIMENTO:
N e *K* indicano, rispettivamente, il numero di oggetti presenti in due buffer comunicanti secondo uno schema produttore-consumatore

$$P_{N,K}(n,k) \hat{=} \Pr\{\overset{\text{probabilità congiunta}}{N = n \cap K = k}\}$$

Ripartizione congiunta

$$n = 0,1,2, \ldots \;\; ; \;\; k = 0,1,2, \ldots \; ;$$

la marginale è pari alla somma dei possibili valori di k

$$P_N^{(m)}(n) \hat{=} \sum_{k=0}^{\infty} P_{N,K}(n,k)$$

Ripartizione marginale

Ripartizione marginale CUMULATIVA

$$C_N^{(m)}(l) = \sum_{n=0}^{l} P_N^{(m)}(n)$$

INDIPENDENZA STOCASTICA

Variabili aleatorie indipendenti <u>se e solo se</u>:

distribuzione congiunta
resta una probabilità
congiunta

$$F_{X_1,X_2}(t_1,t_2)=F_{X_1}^{(m)}(t_1)F_{X_2}^{(m)}(t_2)$$

prodotto tra le marginali
che sono probabilità

saremmo tentati di definire la densità anche come prodotto

$$f_{X_1,X_2}(t_1,t_2)\hat{=}f_{X_1}^{(m)}(t_1)f_{X_2}^{(m)}(t_2)$$

$$0\le t_1<\infty,\quad 0\le t_2<\infty$$

GIUSTIFICAZIONE:

partendo dalla definizione
di densità, ovvero che
la realizzazione sia compresa
tra t e t più un delta, si va a
giustificare l'indipendeza
stocastica

$$\boxed{\begin{aligned}&P(t_1<X_1\le t_1+dt_1\cap t_2<X_2\le t_2+dt_2)\\&\approx f(t_1,t_2)dt_1dt_2\hat{=}f(t_1)dt_1\cdot f(t_2)dt_2\\&=P(t_1<X_1\le t_1+dt_1)\cdot P(t_2<X_2\le t_2+dt_2)\end{aligned}}$$

Distribuzione congiunta e distribuzione marginale

"Spiegazione" non rigorosa di Trivedi della congiunta

Le funzioni di distribuzione marginale si ricavano a partire dalla distribuzione congiunta:

rilasso il vincolo

$$(X_1 \le t_1) \equiv (X_1 \le t_1 \cap X_2 < \infty) = \lim_{t_2 \to \infty} (X_1 \le t_1 \cap X_2 \le t_2)$$

non finisce per forza dopo t_1.
$X_2 < t_1$ è ammesso

$$P(X_1 \le t_1) = P\left(\lim_{t_2 \to \infty} (X_1 \le t_1,\, X_2 \le t_2)\right) = \lim_{t_2 \to \infty} P(X_1 \le t_1,\, X_2 \le t_2)$$

sfruttando la continuità, ovvero una funzione è continua se il valore del limite nel punto coincide con il valore della funzione.

$$F_{X_1}^{(m)}(t_1) = \lim_{t_2 \to \infty} F_{X_1,X_2}(t_1, t_2)$$

Se le variabili aleatorie non sono indipendenti lavoro con la probabilità congiunta

$$F_{X_2}^{(m)}(t_2) = \lim_{t_1 \to \infty} F_{X_1,X_2}(t_1, t_2)$$

Evento limite e sua probabilità

vista all'inizio del corso

$$E_1 \supset E_2 \supset \ldots \supset E_n \supset \ldots \qquad \Rightarrow \lim_{n\to\infty} E_n = \lim_{n\to\infty} \bigcap_{i=1}^{n} E_i \ \hat{=} \supset E$$

$$E_1 \subset E_2 \subset \ldots \subset E_n \subset \ldots \qquad \Rightarrow \lim_{n\to\infty} E_n = \lim_{n\to\infty} \bigcup_{i=1}^{n} E_i \ \hat{=} \subset E$$

$$\boxed{TEOREMA: \quad P(\subset E) \,\hat{=}\, P(\lim_{n\to\infty} E_n) = \lim_{n\to\infty} P(E_n)}$$

SEQUENZA CRESCENTE

PROVA:

$$P(E) \,\hat{=}\, P\left(\lim_{n\to\infty} E_n\right) = P\left(\lim_{n\to\infty} \bigcup_{i=1}^{i=n} E_i\right) = P\left(\bigcup_{i=1}^{\infty} E_i\right)$$

$$= P(E_1) + \sum_{i=1}^{\infty} \left[P(E_{i+1}) - P(E_i)\right]$$

$$= P(E_1) + \lim_{n\to\infty} \sum_{i=1}^{n} \left[P(E_{i+1}) - P(E_i)\right]$$

$$= \lim_{n\to\infty} P(E_n)$$

CONTINUITA' DI ASSEGNAMENTO PROBABILISTICO

Esempio classico: Il problema dell'incontro (1/4)

problema di sincronia

I signori Tizio e Caio si trattengono tutte le sere in un noto bar di Rende, rispettivamente Δt_1 minuti e Δt_2 minuti, nell'intervallo [0,T], sotto le ipotesi seguenti: $0 <= \Delta t_1 <= T$, e $0 <= \Delta t_2 <= T$.

sono diversi i tempi di attesa che rientrano nell'intervallo

Escludiamo che i due abbiano un appuntamento al bar e siano pure puntuali, ma assumiamo che gli istanti di arrivo, x e y, di entrambi siano il risultato di un comportamento "congiuntamente aleatorio".

Più precisamente, col termine "congiuntamente aleatorio" intendiamo sintetizzare un comportamento non necessariamente indipendente fra i due e neppure completamente casuale per quanto riguarda l'istante di arrivo al bar di ciascuno dei due.

Io e Giulio tutte le sere alle 19 finiamo e ci dirigiamo al bar dell'università. Senza darci un appuntamento se ci incontriamo al bar ci fermiamo per un aperol. Io aspetto t1 minuti quando arrivo mentre Giulio aspetta t2 minuti. Quindi può capitare che magari una sera io termino alle 18:50 e quindi mi trattengo di meno al bar oppure finisco più tardi e passo quando Giulio se nè già andato via.

Esempio classico: Il problema dell'incontro (2/4)

Usando lo spazio reale euclideo a due dimensioni, vogliamo individuare i sottospazi corrispondenti agli eventi di interesse seguenti:

C= "Tizio arriva non dopo Caio" e D= "Tizio e Caio si incontrano"

Indicando con x e y i rispettivi istanti di arrivo, risulta:

Tizio arriva prima di Caio

$$C \hat{=} \left\{ (x, y) \in [0, T] \times [0, T] : x \leq y \right\}$$

y arriva entro il tempo tollerato da x

$$D \hat{=} D_1 \cup D_2$$

$$\text{con } D_1 \hat{=} \left\{ x, y \in [0, T] : x \leq y \cap y \leq x + \Delta t_1 \right\}$$

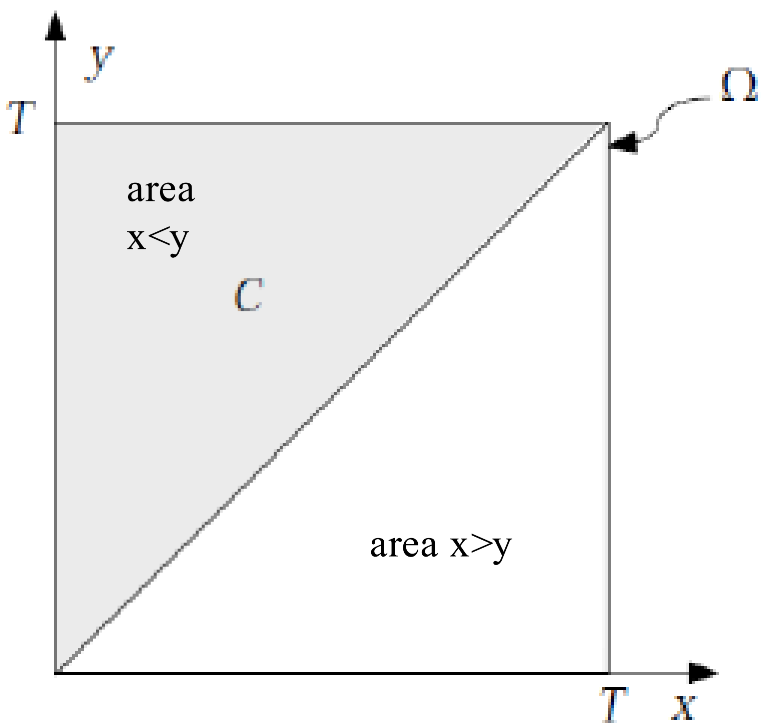
$$\text{e } D_2 \hat{=} \left\{ x, y \in [0, T] : y \leq x \cap x \leq y + \Delta t_2 \right\}$$

Può avvenire anche il contrario. Da notare che non aspettano la stessa quantità di tempo Δt_i

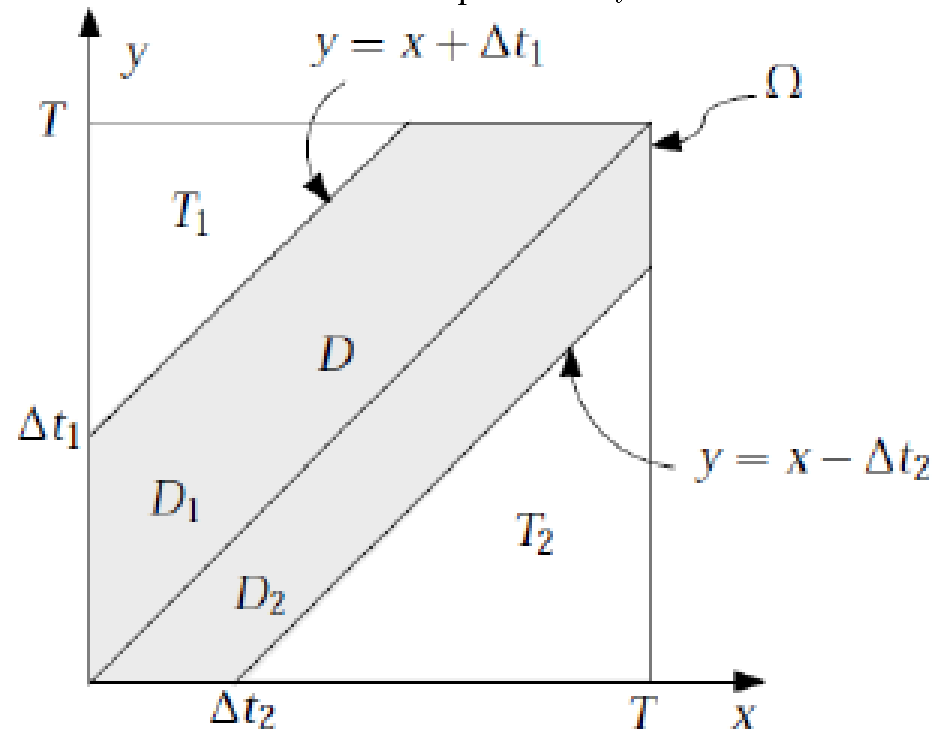
Esempio classico: Il problema dell'incontro (3/4)

RAPPRESENTAZIONE DEGLI EVENTI DI INTERESSE:

sulla diagonale arrivano allo stesso istante



aree diverse perché Δt_i diversi



Avendo una funzione densità che attribuisca probabilità ai risultati elementari, potremmo calcolare la probabilità degli eventi di interesse per integrazione prima su C e poi su D.

Esempio classico: Il problema dell'incontro (4/4)

CALCOLO DELLE PROBABILITA' DEGLI EVENTI DI INTERESSE:

Rinunciando alla densità congiunta, ma con le seguenti ipotesi :

1) Tizio e Caio sono indipendenti;
2) L'istante di arrivo di ciascuno di loro è la realizzazione di una variabile aleatoria con distribuzione uniforme in [0,T].

RISULTATO DEL CALCOLO NELLE IPOTESI 1) E 2):

$x<y$

Prob("Tizio arriva prima di Caio") =

$= \text{AREA\_C} / \text{AREA\_}\Omega = (1/2)T^2 / T^2 = 1/2$ appunto area divisa a metà dalla diagonale nella slide precedente

Prob("Tizio e Caio si incontrano") $= \text{AREA\_D} / \text{AREA\_}\Omega =$

$= (\text{AREA\_}\Omega - \text{AREA\_T}_1 - \text{AREA\_T}_2) / \text{AREA\_}\Omega$

$= T^2 - ((T-\Delta t_1)^2/2) - ((T-\Delta t_2)^2/2) / T^2$

(Con $T= 60$min e $\Delta T_1= \Delta T_2 =5$min allora Prob("incontro") = 0.139)

CDF=densità
PDF=distribuzione
partono assieme

Si cerca il ritardo di sincronia

Example 3.29. Romeo and Juliet have a date at a given time, and each, independently, will be late by an amount of time that is exponentially distributed with parameter λ. What is the PDF of the difference between their times of arrival?

Let us denote by X and Y the amounts by which Romeo and Juliet are late, respectively. We want to find the PDF of $Z = X - Y$, assuming that X and Y are independent and exponentially distributed with parameter λ. We will first calculate the CDF $F_Z(z)$ by considering separately the cases $z \geq 0$ and $z < 0$ (see Fig. 3.26).

X arriva dopo Y arriva dopo

nel esempio non dice che partono allo stesso istante temporale, perchè si usano modelli esponenziali quindi c'è l'assenza di memoria

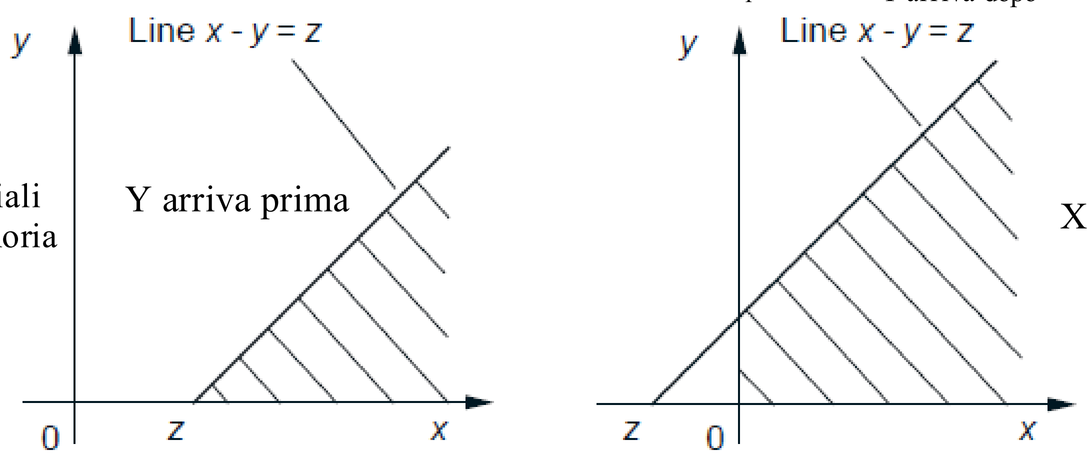


Figure 3.26: The calculation of the CDF of $Z = X - Y$ in Example 3.29. To obtain the value $\mathbf{P}(X - Y > z)$ we must integrate the joint PDF $f_{X,Y}(x, y)$ over the shaded area in the above figures, which correspond to $z \geq 0$ (left side) and $z < 0$ (right side).

Combining the two cases $z \geq 0$ and $z < 0$, we obtain

$$F_Z(z) = \begin{cases} 1 - \dfrac{1}{2}e^{-\lambda z} & \text{if } z \geq 0, \\ \dfrac{1}{2}e^{\lambda z} & \text{if } z < 0, \end{cases}$$

We now calculate the PDF of Z by differentiating its CDF. We obtain

$$f_Z(z) = \begin{cases} \dfrac{\lambda}{2}e^{-\lambda z} & \text{if } z \geq 0, \\ \\ \dfrac{\lambda}{2}e^{\lambda z} & \text{if } z < 0, \end{cases}$$

or

$$f_Z(z) = \frac{\lambda}{2}e^{-\lambda|z|}.$$

This is known as a **two-sided exponential PDF**, also known as the **Laplace PDF**.

For $z \geq 0$, we have (see the left side of Fig. 3.26)

$$
\begin{aligned}
F_Z(z) &= \mathbf{P}(X - Y \leq z) \\
&= 1 - \mathbf{P}(X - Y > z) \;= 1 - \int_0^\infty \left(\int_{z+y}^\infty f_{X,Y}(x,y)\,dx \right) dy \\
&= 1 - \int_0^\infty \lambda e^{-\lambda y} \left(\int_{z+y}^\infty \lambda e^{-\lambda x}\,dx \right) dy \;= 1 - \int_0^\infty \lambda e^{-\lambda y} e^{-\lambda(z+y)}\,dy \\
&= 1 - e^{-\lambda z} \int_0^\infty \lambda e^{-2\lambda y}\,dy \;= 1 - \frac{1}{2} e^{-\lambda z}.
\end{aligned}
$$

For the case $z < 0$, we can use a similar calculation, but we can also argue using symmetry. Indeed, the symmetry of the situation implies that the random variables $Z = X - Y$ and $-Z = Y - X$ have the same distribution. We have

$$
F_Z(z) = \mathbf{P}(Z \leq z) = \mathbf{P}(-Z \geq -z) = \mathbf{P}(Z \geq -z) = 1 - F_Z(-z).
$$

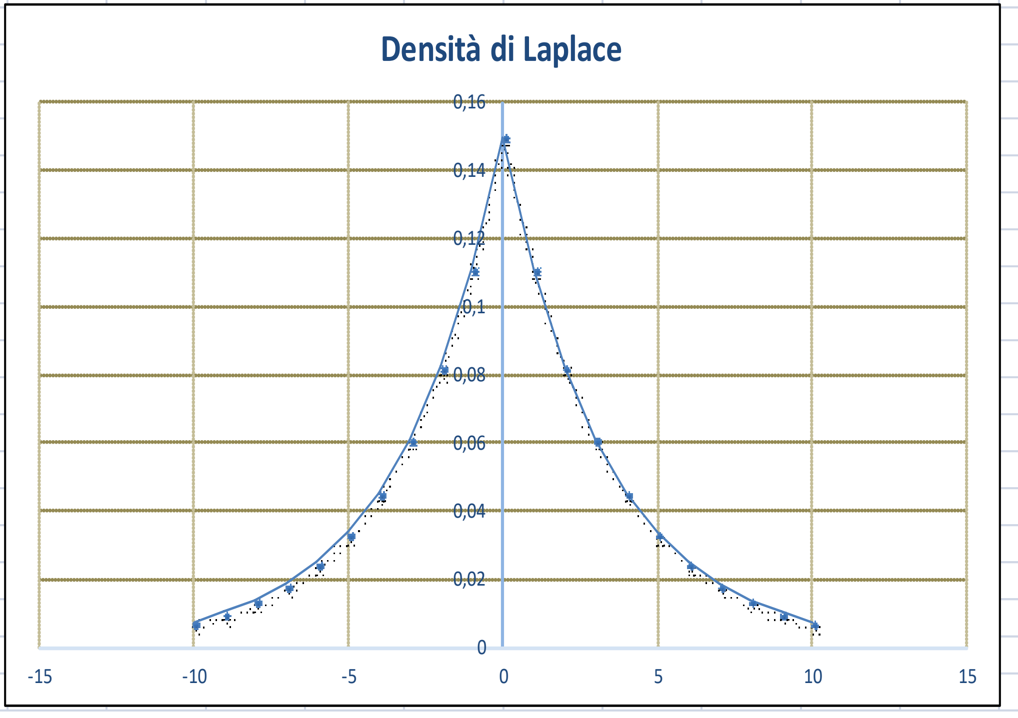
With $z < 0$, we have $-z \geq 0$ and using the formula derived earlier,

$$
F_Z(z) = 1 - F_Z(-z) = 1 - \left(1 - \frac{1}{2} e^{-\lambda(-z)}\right) = \frac{1}{2} e^{\lambda z}.
$$

Densità di Laplace con EXCEL

Costituita da due esponenziali come visto ed impiegata per problemi di ritardo di sincronia.

z	f(z)
-10	0,007468
-9	0,010081
-8	0,013608
-7	0,018368
-6	0,024795
-5	0,03347
-4	0,045179
-3	0,060985
-2	0,082322
-1	0,111123
0	0,15
1	0,111123
2	0,082322
3	0,060985
4	0,045179
5	0,03347
6	0,024795
7	0,018368
8	0,013608
9	0,010081
10	0,007468



La distribuzione congiunta in Affidabilità

(nei modelli di sopravvivenza)

$$F_{X_1,X_2}(t,t) \hat{=} P(X_1 \leq t \cap X_2 \leq t) \qquad 0 \leq t < \infty$$

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 NON INTERESSA LA DISPOSIZIONE SE IN SERIE O PARALLELO PERCHÈ HO CAUSE DI GUASTO INDIVIDUALI!

AFFIDABILITA' CONGIUNTA DEI DUE COMPONENTI:
(sopravvivenza)

$$R(t,t) \hat{=} P(X_1 > t \cap X_2 > t)$$

t è il tempo del PRIMO guasto, ma non si sa quale dei due componenti si guasti per primo!

Modello primitivo quando non è possibile caratterizzare separatamente il tempo al guasto di uno qualunque dei due componenti, indipendentemente dall'altro.

Nel caso ad una dimensione R(t)=1-F(t). Vale anche in questo caso? NO

La distribuzione congiunta in Affidabilità

$$R(\bar{t},\bar{t}) \triangleq P\big((X_1 \cap X_2) > \bar{t}\big) \triangleq \int_{u_1=\bar{t}}^{\infty}\int_{u_2=\bar{t}}^{\infty} f_{X_1,X_2}(u_1,u_2)\,du_1 du_2$$

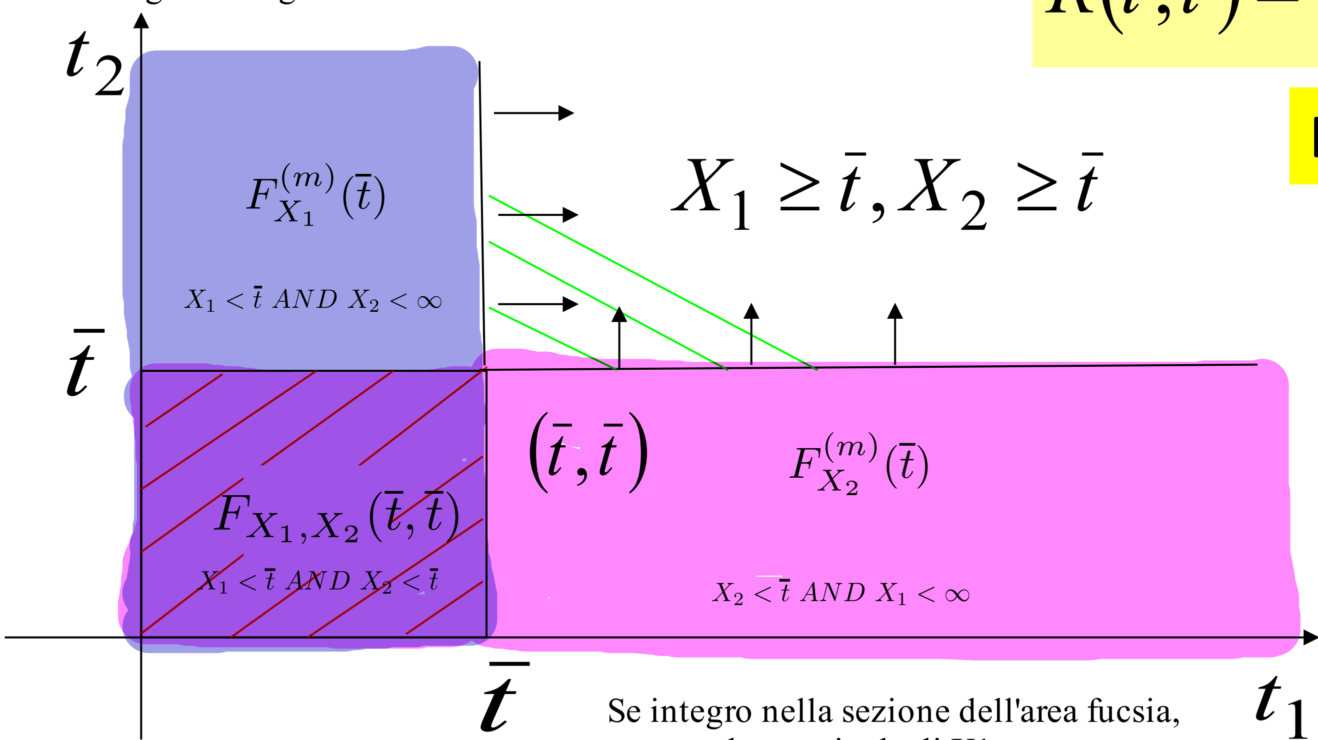
Il complemento a entrambi che tornano
prima di t, non è tutti e due non tornano dopo di t,
ma che uno dei due torni prima di t

Importante

$$R(\bar{t},\bar{t}) = 1 - F_{X_1,X_2}(\bar{t},\bar{t})$$

E' FALSO!

Se integro nella sezione dell'area blu,
ottengo la marginale di X2



Se integro nella sezione dell'area fucsia,
ottengo la marginale di X1

Distribuzione congiunta e distribuzione marginale

Viaggi, incontri e ritardi di sincronia a parte,

QUANDO ANCORA SI USA UNA DISTRIBUZIONE CONGIUNTA?

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 (degenze ripetute in ospedale, fidanzamento e matrimonio);

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 (attesa dal parrucchiere e numero di persone trovate al proprio arrivo);

Distribuzione della somma di variabili aleatorie (indipendenti)

Si vuole calcolare la funzione di distribuzione di una somma di variabili aleatorie, ovvero di una funzione di variabili aleatorie

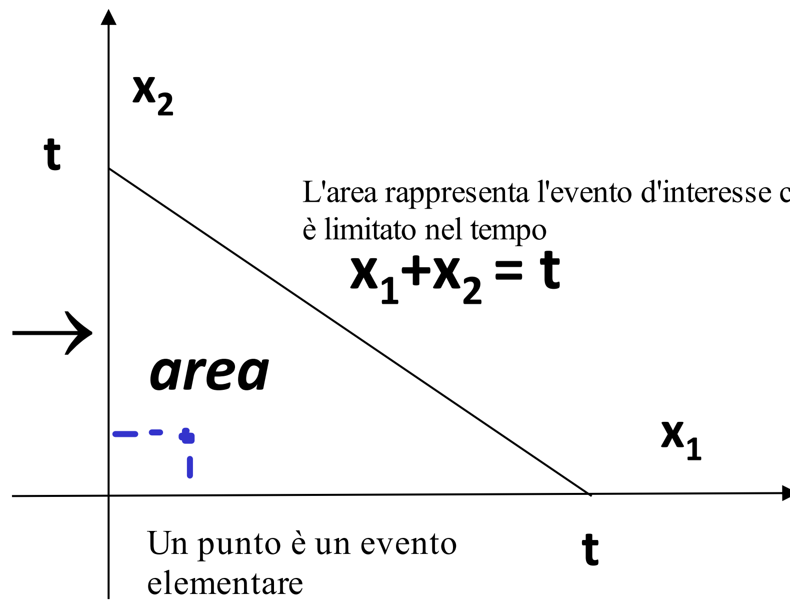
L'ipotesi di indipendeza nascerà più avanti

$$\{(x_1, x_2);\ x_1 \geq 0,\ x_2 \geq 0\}$$

v.a. tempo che passa dall'inizio della prima attività alla fine della seconda attività

$$X \hat{=} X_1 + X_2 \ \Rightarrow\ X \leq t \ \rightarrow$$

Durata complessiva evento d'interesse



$$F_X(t) \hat{=} \Pr\{X_1 + X_2 \leq t\}$$

Per ipotesi le variabili sono ancora dipendenti tra loro quindi utilizzo la congiunta

$$F_X(t) = \iint_{area} f_{X_1 X_2}$$

$$F_X(t) = \int_{x_2=0}^{t} \int_{x_1=0}^{t-x_2} f_{X_1,X_2}(x_1, x_2)\, dx_1 dx_2$$

$$\equiv \int_{x_1=0}^{t} \int_{x_2=0}^{t-x_1} f_{X_1,X_2}(x_1, x_2)\, dx_1 dx_2$$

SENZA INDIPENDENZA E' LA FORMULA FINALE!

Distribuzione della somma di variabili aleatorie (indipendenti)

Mi gioco l'ipotesi di indipendenza,
quindi passo a prodotto di densità individuali come visto nella lezione precedente

$$\int_{x_1=0}^{t}\int_{x_2=0}^{t-x_1} f_{X_1,X_2}(x_1,x_2)dx_2 dx_1 = \int_{x_1=0}^{t}\left(\int_{x_2=0}^{t-x_1} f_{X_2}(x_2)dx_2\right) f_{X_1}(x_1)dx_1$$

probabilità che il componente attivo si guasti entro t e che il
si guasti entro t-x1

In verde probabilità che X2 termini prima di t-x1. In arancione che il primo componente si guasti entro x1
Non è escluso che x1 duri t-epsilon o che duri 0+epsilon

$$= \int_{x_1=0}^{t} F_{X_2}(t-x_1) f_{X_1}(x_1)dx_1 =$$

integrale di convoluzione
La convoluzione nasce dall'indipendenza

oppure se il primo integrale è tra $x_1 = 0, t - x_2$

$$= \int_{x_2=0}^{t} F_{X_1}(t-x_2) f_{X_2}(x_2)dx_2$$

***X1* e *X2* variabili aleatorie <u>*indipendenti*</u> !**

Se conoscessi l'istante in cui X1 si guasta
l'integrale non sarebbe necessario, basterebbe
$F_{X_2}(t-\overline{x}_1)f_{X_1}(\overline{x}_1)dx_1$

<u>*convoluzione* delle funzioni densità dell'una con la distribuzione dell'altra</u>

$$\int_{x_1=0}^{t} F_{X_2}(t-x_1 \mid x_1)\cdot f_{X_1}(x_1)dx_1$$

sarebbe una probabilità totale

La sostituzione $F_{X_2}(t-x_1) = F_{X_2}(t-x_1|x_1)$
vale perchè si ha indipendeza

Convoluzione di esponenziali identiche e indipendenti

Sommando n esponenziali

$$F_{X_i}(t) = 1 - e^{-\lambda t} \quad i = 1,2,\dots,n \quad \rightarrow \quad X \triangleq X_1 + X_2 + \dots + X_n$$

TESI:

$$f_X(t) = \frac{(\lambda t)^{n-1}}{(n-1)!} \lambda e^{-\lambda t}$$

DENSITA'
DI ERLANG DI ORDINE "n"

PROVA: per induzione

$$\text{Posto}: \; Y \triangleq X_1 + X_2 + \dots + X_{n-1} \quad \Rightarrow \quad X = Y + X_n$$

$$\boxed{\Rightarrow f_X(t) = \frac{d}{dt} F_X(t) = \frac{d}{dt} \int_{u=0}^{t} F_{X_n}(t-u) \cdot f_Y(u) \cdot du}$$

derivando

$$= \int_{u=0}^{t} f_{X_n}(t-u) f_Y(u) du = \int_{u=0}^{t} f_{X_n}(t-u) \cdot \frac{(\lambda u)^{n-2}}{(n-2)!} e^{-\lambda u} \lambda \cdot du$$

esplicito in Erlang

sostituzione di u

$$= \int_{u=0}^{t} \lambda e^{-\lambda(t-u)} \cdot e^{-\lambda u} \frac{(\lambda u)^{n-2}}{(n-2)!} d(\lambda u) = \lambda e^{-\lambda t} \frac{(\lambda t)^{n-1}}{(n-1)!}$$

Convoluzione di esponenziali identiche

non è domanda

$$F_X(t) = \int_0^t e^{-\lambda u} \frac{(\lambda u)^{n-1}}{(n-1)!} d(\lambda u)$$

Si risolve per parti

$$= e^{-\lambda t} \frac{(\lambda t)^n}{n!} + \int_0^t e^{-\lambda u} \frac{(\lambda u)^n}{n!} d(\lambda u) = e^{-\lambda t} \frac{(\lambda t)^n}{n!} + \int_0^t e^{-\lambda u} d\left(\frac{(\lambda u)^{n+1}}{(n+1)!} \right)$$

$$= e^{-\lambda t} \frac{(\lambda t)^n}{n!} + e^{-\lambda t} \frac{(\lambda t)^{n+1}}{(n+1)!} + \int_0^t e^{-\lambda u} d\left(\frac{(\lambda u)^{n+2}}{(n+2)!} \right) = \ldots$$

$$\ldots = e^{-\lambda t} \sum_{i=n}^{\infty} \frac{(\lambda t)^i}{i!}$$

DISTRIBUZIONE DI ERLANG DI ORDINE “n”

Dopo tutto quanto ottengo

Il modello di Erlang -n

Densità:

$$\boxed{f_X(t) = \frac{(\lambda t)^{n-1}}{(n-1)!}\lambda e^{-\lambda t}}$$

Distribuzione:

$$F_X(t) = \int_0^t e^{-\lambda u}\frac{(\lambda u)^{n-1}}{(n-1)!}d(\lambda u) \qquad \ldots = e^{-\lambda t}\sum_{i=n}^{\infty}\frac{(\lambda t)^i}{i!}$$

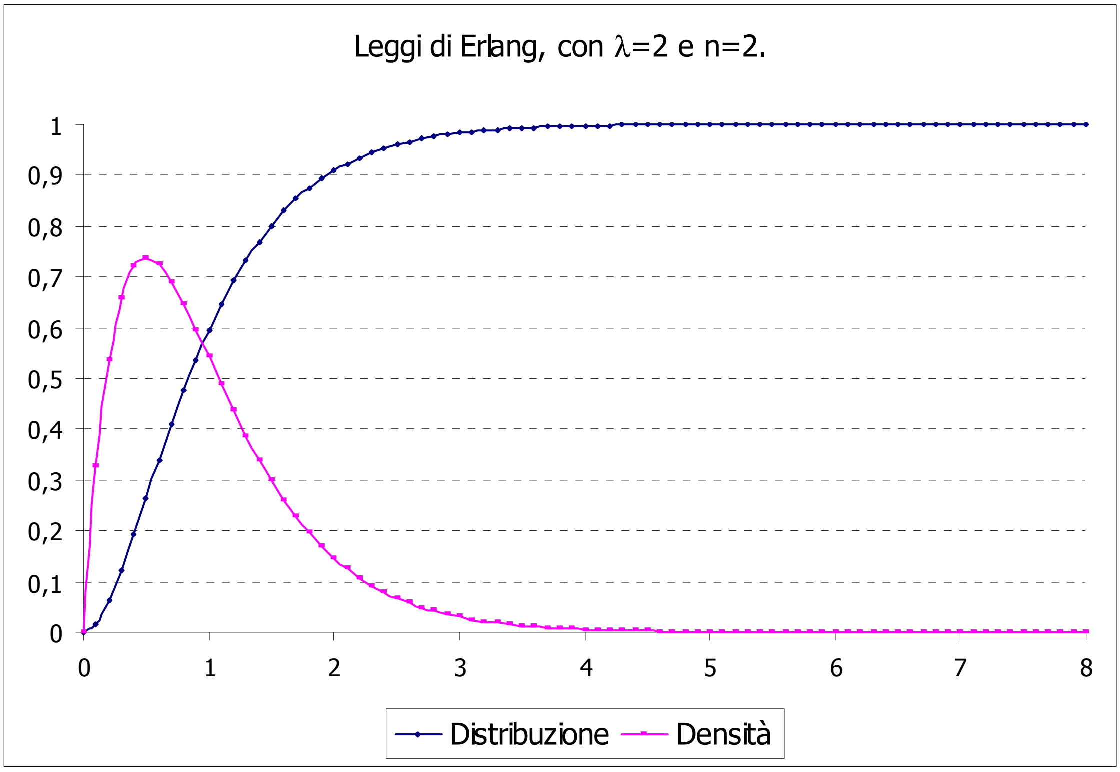
$$= e^{-\lambda t}\left(\sum_{i=0}^{\infty}\frac{(\lambda t)^i}{i!} - \sum_{i=0}^{n-1}\frac{(\lambda t)^i}{i!}\right) = e^{-\lambda t}\left(e^{\lambda t} - \sum_{i=0}^{n-1}\frac{(\lambda t)^i}{i!}\right)$$

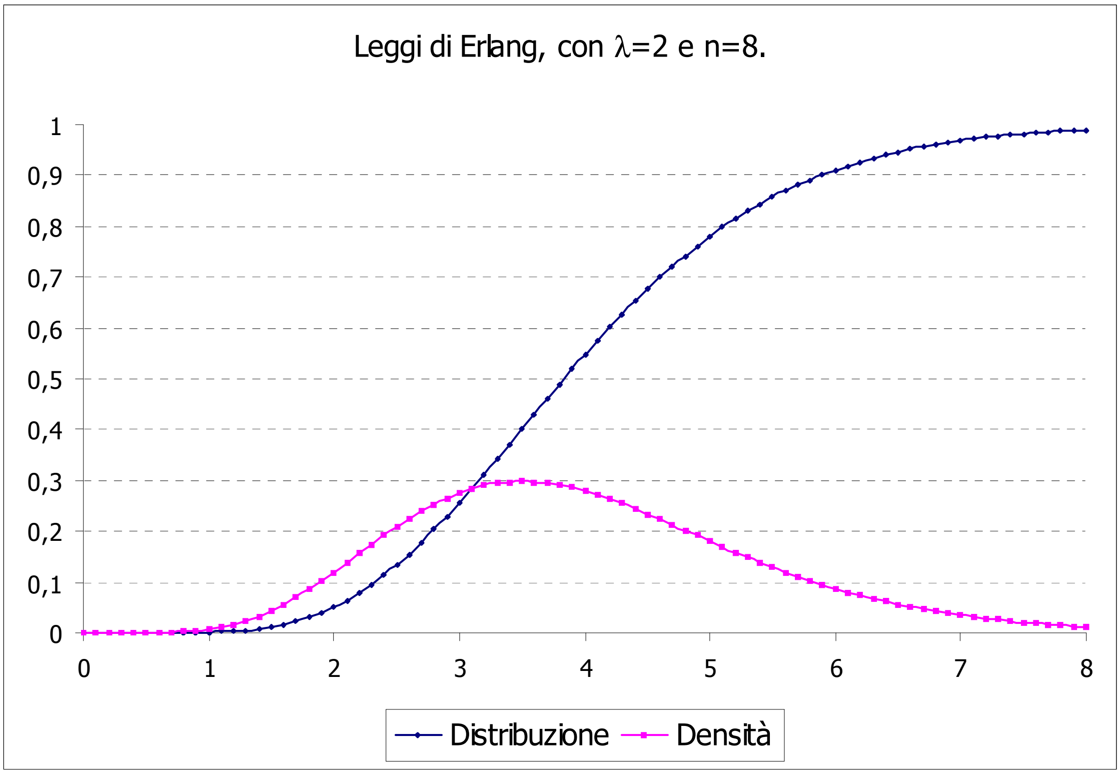
ci interessa fermarci a n-1

$$\boxed{= 1 - e^{-\lambda t}\sum_{i=0}^{n-1}\frac{(\lambda t)^i}{i!}}$$

quindi questa è la distribuzione che si usa per fare calcoli
Erlang non ha assenza di memoria perché mi dice la posizione in cui mi trovo, grazie all'indice!

Illustrazione del Modello di Erlang (1)





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 scegliamo 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:

$$X \hat{=} X_1 + X_2 + \ldots + X_n \quad \text{con} \quad X_i \approx exp(\lambda) \quad i = 1,\ldots,n$$ (anche <u>indipendenti</u>)

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:

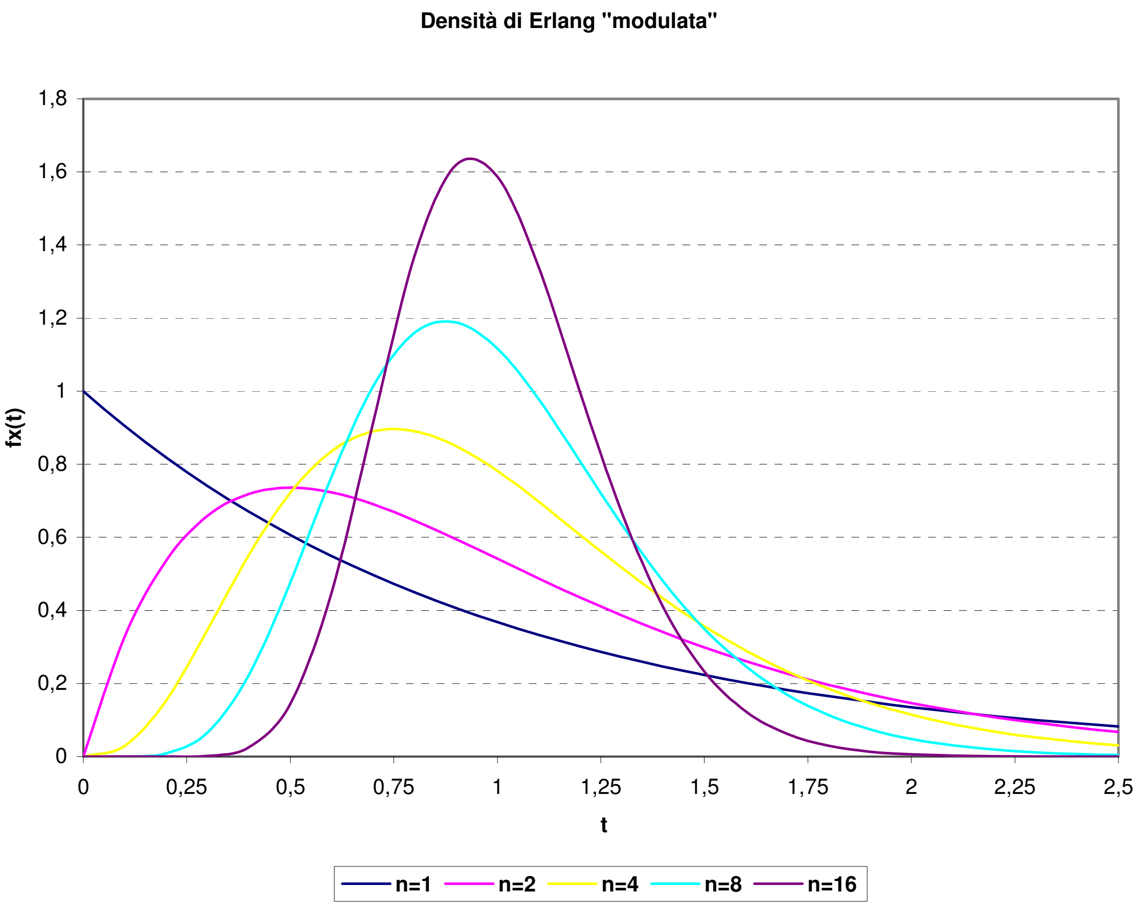
$$\lambda \hat{=}\ n\,\frac{1}{E[X]}$$

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 <u>Legge di Erlang MODULATA</u>.



Esempio

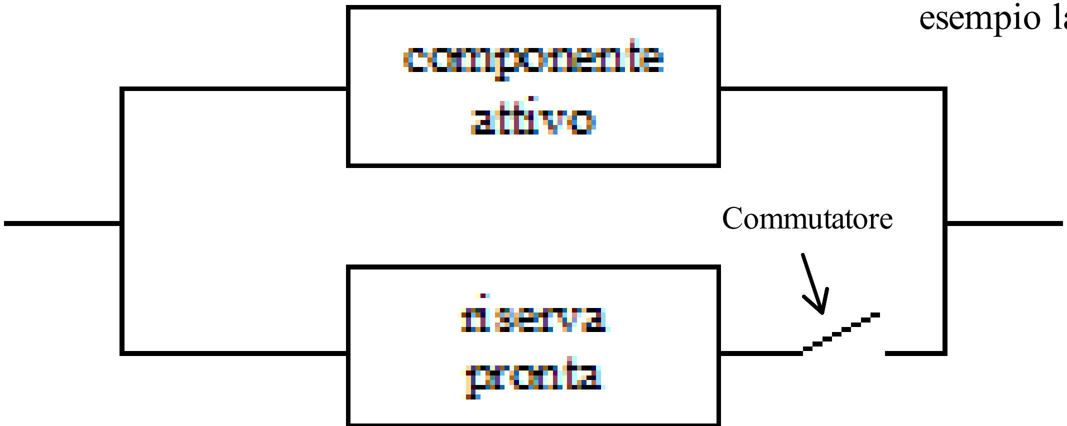
Si supponga che un componente, una volta guasto, possa essere riparato secondo un processo di riparazione che si compone di quattro fasi in sequenza. La durata media dell'intero processo è stimata in 2 unità di tempo e non si hanno informazioni sulle durate parziali medie delle singole fasi: potrebbero 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?

Affidabilità dei sistemi con riserve pronte

Componenti NON identici quindi affidabilità diversa. Ad esempio la riserva è più scadente per risparmiare soldi



X = "Vita del sistema";

X 1 = "Vita comp. att.";

X 2 = "Vita riserva att.";

La commutazione è il passaggio da componente attivo a componente di riserva

- la commutazione sia istantanea e che riesca con certezza = PERFETTA
- la riserva non possa subire guasti quando permane pronta.

$$\Rightarrow X \triangleq X_1 + X_2$$

$$X_i \approx \exp(\lambda_i) \quad i = 1, 2.$$

tempo all'ultimo guasto

$\lambda_1 \neq \lambda_2$ qualità diversa ovvero affidabilità diversa

$$F_X(t) = \int_{x_1=0}^{t} \left[1 - e^{-\lambda_2 (t - x1)}\right] \lambda_1 e^{-\lambda_1 x_1} dx_1$$

$$= 1 - \left(\frac{\lambda_2}{\lambda_2 - \lambda_1} e^{-\lambda_1 t} + \frac{\lambda_1}{\lambda_1 - \lambda_2} e^{-\lambda_2 t} \right)$$

$$R(t) = 1 - F_X(t)$$

Affidabilità dei sistemi con riserve pronte

GENERALIZZAZIONE:
"n-1" componenti di riserva al componente attivo. Tutti identici.

$$X \hat{=} X_1 + X_2 + \ldots + X_n \qquad X_i \approx exp(\lambda) \quad i = 1,\ldots,n$$

$$\boxed{R(t) = e^{-\lambda t} \sum_{i=0}^{n-1} \frac{(\lambda t)^i}{i!}}$$

ottenuto da $R(t) = 1 - F_X(t)$

Affidabilità nel

MODELLO DI ERLANG

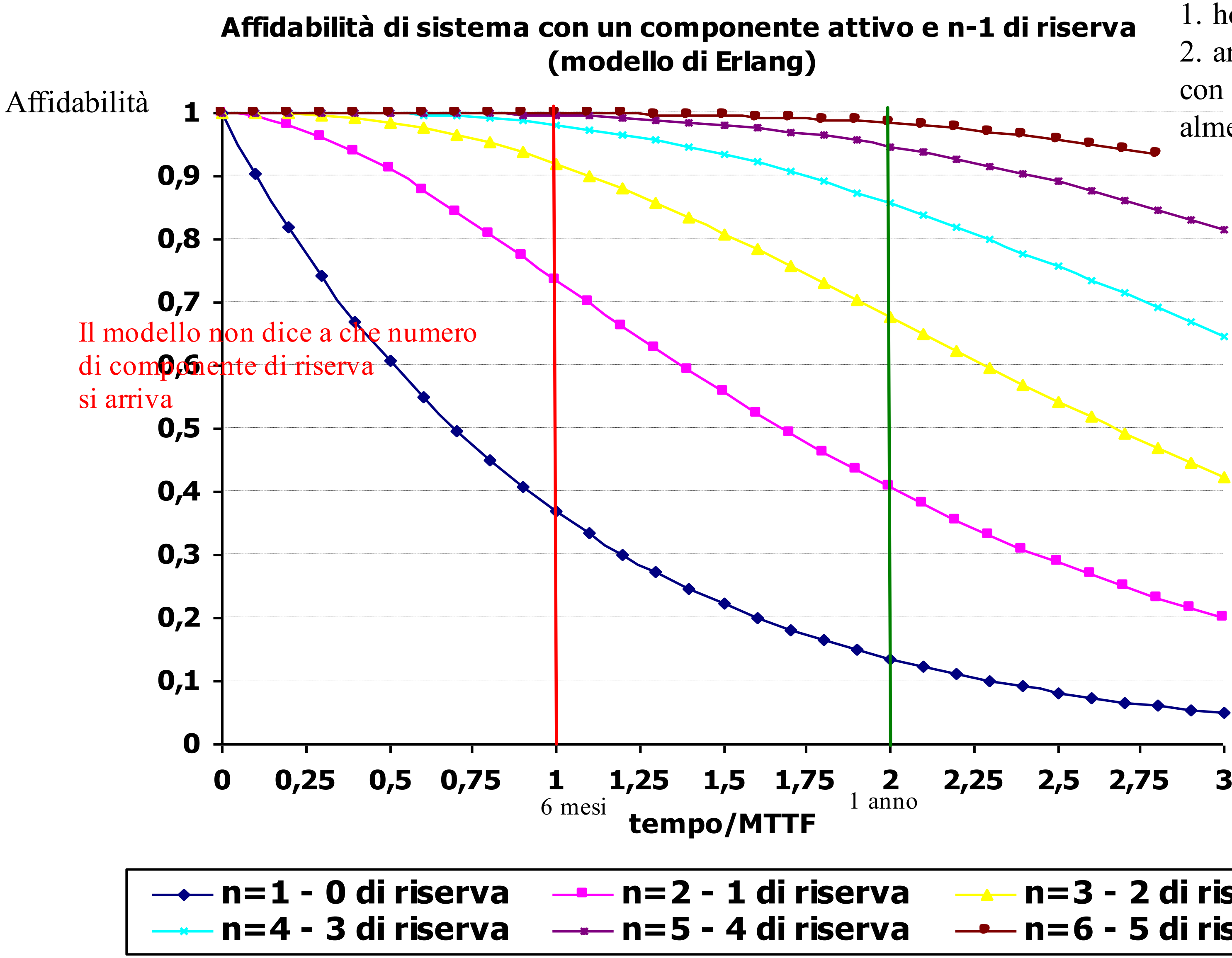
Ipotesi:

- *la commutazione riesce sicuramente;*
- *la commutazione si realizza in un tempo trascurabile ;*
- *i componenti sono identici e non riparabili;*
- *la legge comune di guasto è esponenziale di parametro λ.*

Affidabilità dei sistemi con riserve pronte

Fissato un istante di tempo calcolo la probabilità che:

1. ho rotto tutti i componenti
2. arrivo all'istante fissato con probabilità> 0.9 con almeno una riserva



Fisso l'istante di tempo a 1 che corrisponde a 6 mesi ad esempio. Allora per ogni curva posso calcolare la probabilità di che tutti i componenti siano rotti come 1-affidabilità%, ad esempio per la blu ho 1-0.38=0.62. Mentre se volessi superare con 0.9 di probabilità i 6 mesi dovrei adoperare un sistema con almeno 3-2 di riserva, se invece volessi superare un anno dovrei adoperare un sistema 5-4 di riserva.

Il modello non dice a quale di componente di riserva si arriva, anche perchè non comunica quando si rompono i componenti!

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 è la condizionata
X è la condizionante

$$"Y \leq y / X = x" \hat{=} \lim_{\Delta x \to 0} "Y \leq y / x \leq X \leq x + \Delta x"$$

e assumere che valga la seguente:

$$P\left(\lim_{\Delta x \to 0} "Y \leq y / x \leq X \leq x + \Delta x"\right) = \lim_{\Delta x \to 0} P(Y \leq y / x \leq X \leq x + \Delta x)$$

Cosicché :

$$F_{Y/X}(y/x) \hat{=} \lim_{\Delta x \to 0} P(Y \leq y / x \leq X \leq x + \Delta x)$$

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:

$$F_{Y|X}(y \mid x) = \int_{v=0}^{y} \frac{f_{Y,X}(v,x)}{f_X^{(m)}(x)} \, dv$$

Da cui è immediato riconoscere la funzione densità condizionata:

$$f_{Y/X}(v/x) \hat{=} \frac{f_{Y,X}(x,v)}{f_X^{(m)}(x)}$$

Dalla formula di calcolo è possibile ricavare la densità marginale per la variabile aleatoria Y

$$F_Y^{(m)}(y) = \int_{x=0}^{\infty} F_{Y|X}(y|x) \cdot f_X^{(m)}(x) dx$$

dove ad esempio se ha l'attesa di un client Y e la durata del servizio X, l'integrale viene calcolato per tutti i possibili valori di durata del servizio. Inoltre la margina di X rappresenta la probabilità che il servizio arrivi all'istante x e finisca subito dopo entro un dx. Infine sorge una domanda: il tempo di attesa può essere calcolato come un modello primitivo? È linearizzabile? La risposta in entrambi i casi è no

Derivazione della formula di calcolo

Prima di ricavarla, è bene far vedere che la probabilità dell'evento condizionante è misurabile con la densità marginale:

$$P(x \le X \le x+\Delta x) = \lim_{y\to\infty} P(x \le X \le x+\Delta x \cap Y \le y)$$

$$= \int_{u=x}^{x+\Delta x} \int_{v=0}^{\infty} f_{Y,X}(v,u)dvdu \;\; = f_X^{(m)}(x)\Delta x$$

A questo punto:

Formula concettuale della probabilità condizionata per Trivedi

$$F_{Y/X}(y/x) \triangleq \lim_{\Delta x\to 0} P(Y \le y / x \le X \le x+\Delta x)$$

$$= \lim_{\Delta x\to 0} \frac{P(Y \le y \cap x \le X \le x+\Delta x)}{\lim_{y\to\infty} P(x \le X \le x+\Delta x \cap Y \le \infty)}$$

$$= \lim_{\Delta x\to 0} \frac{\int_{v=0}^{y} \int_{u=x}^{x+\Delta x} f_{Y,X}(v,u)dudv}{\Delta x \; f_X^{(m)}(x)}$$

$$= \lim_{\Delta x\to 0} \frac{\int_{v=0}^{y} f_{Y,X}(v,\bar{u})\Delta x \; dv}{\Delta x \; f_X^{(m)}(x)} \qquad ove \qquad x \le \bar{u} \le x+\Delta x$$

$$= \int_{v=0}^{y} \frac{f_{Y,X}(v,x)}{f_X^{(m)}(x)} \, dv$$

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: $F_Y^{(m)}(y) = \lim_{x\to\infty} F_{Y,X}(y,x)$

Ricordando che:

$$F_{Y,X}(y,x) \triangleq \int_{v=0}^{y}\int_{u=0}^{x} f_{Y,X}(v,u)dudv$$

si ha:

densità congiunta

da notare come l'integrale rispetto x cambi estremo superiore

$$F_Y^{(m)}(y) = \int_{v=0}^{y}\int_{x=0}^{\infty} f_{Y,X}(v,x)dxdv = \int_{v=0}^{y}[\int_{x=0}^{\infty} f_{Y|X}(v|x)f_X^{(m)}(x)dx]dv$$ cambio l'ordine

$$= \int_{x=0}^{\infty}[\int_{v=0}^{y} f_{Y|X}(v|x)dv]\cdot f_X^{(m)}(x)dx$$

e da qui la <u>formula finale</u>: $F_Y^{(m)}(y) = \int_{x=0}^{\infty} F_{Y|X}(y|x)f_X^{(m)}(x)dx$

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 $Y \leq y \mid X \in (x, x+dx)$ e la probabilità del rispettivo condizionante $X \in (x, x+dx)$.

Infine, per derivazione della formula finale, si ottiene la <u>formula della densità totale</u>:

$$f_Y^{(m)}(y) = \int_{x=0}^{\infty} f_{Y|X}(y|x)f_X^{(m)}(x)dx$$

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:

$$F_Y^{(m)}(y) = \sum_{x=0}^{\infty} F_Y(y \mid x)\cdot P_X^{(m)}(x) \quad \text{e} \quad f_Y^{(m)}(y) = \sum_{x=0}^{\infty} f_Y(y \mid x)\cdot P_X^{(m)}(x).$$

Commutazione imperfetta ed affidabilità

Utilizziamo la condizionante

Un componente attivo ed uno di riserva (identici)

Ho solo una chance per commutare i componenti

v.a. discreta condizionante

$$X = \begin{cases} 1 \; \hat{=} \; commutazione \quad riesce \\ 0 \; \hat{=} \; commutazione \quad fallisce \end{cases}$$

Ipotesi numero 1 si ha un

per semplicità

Meccanismo bernoulliano

per la commutazione

$$\boxed{P_X(1) = c, \quad P_X(0) = 1 - c}$$

c è detto fattore di coverage

Si cerca la distribuzione del tempo di vita del sistema dove

Y = "tempo di vita del sistema"

Ipotesi numero 2:

Leggi di guasto esponenziali

perchè si usa la convoluzione quindi userò Erlang

Passo 1, uso la distribuzione totale

0) $$\boxed{F_{Y|X=0}(t) = 1 - e^{-\lambda t};}$$

esprime il vincolo che la commutazione non riesce, ovvero il sistema dura fino a quando il componente attivo dura

1) $$\boxed{F_{Y|X=1}(t) = 1 - e^{-\lambda t} \sum_{i=0}^{2-1} \frac{(\lambda t)^i}{i!} = 1 - e^{-\lambda t}(1 + \lambda t)}$$

commutazione riuscita

uso Erlang

La 0) è la mortalità, mentre la 1) è l'affidabilità

Distribuzione totale:

non si usa l'integrale!

UN SOLO TENTATIVO!

$$F_Y^{(m)}(t) = F_{Y|X=1}(t) \cdot P_X(1) + F_{Y|X=0}(t) \cdot P_X(0)$$

$$= \left(1 - e^{-\lambda t} - \lambda t e^{-\lambda t}\right) \cdot c + (1 - e^{-\lambda t}) \cdot (1 - c)$$

$$= 1 - (1 + c \cdot \lambda t) \cdot e^{-\lambda t} \qquad \boxed{\Rightarrow R(t) = (1 + c \cdot \lambda t) \cdot e^{-\lambda t}}$$

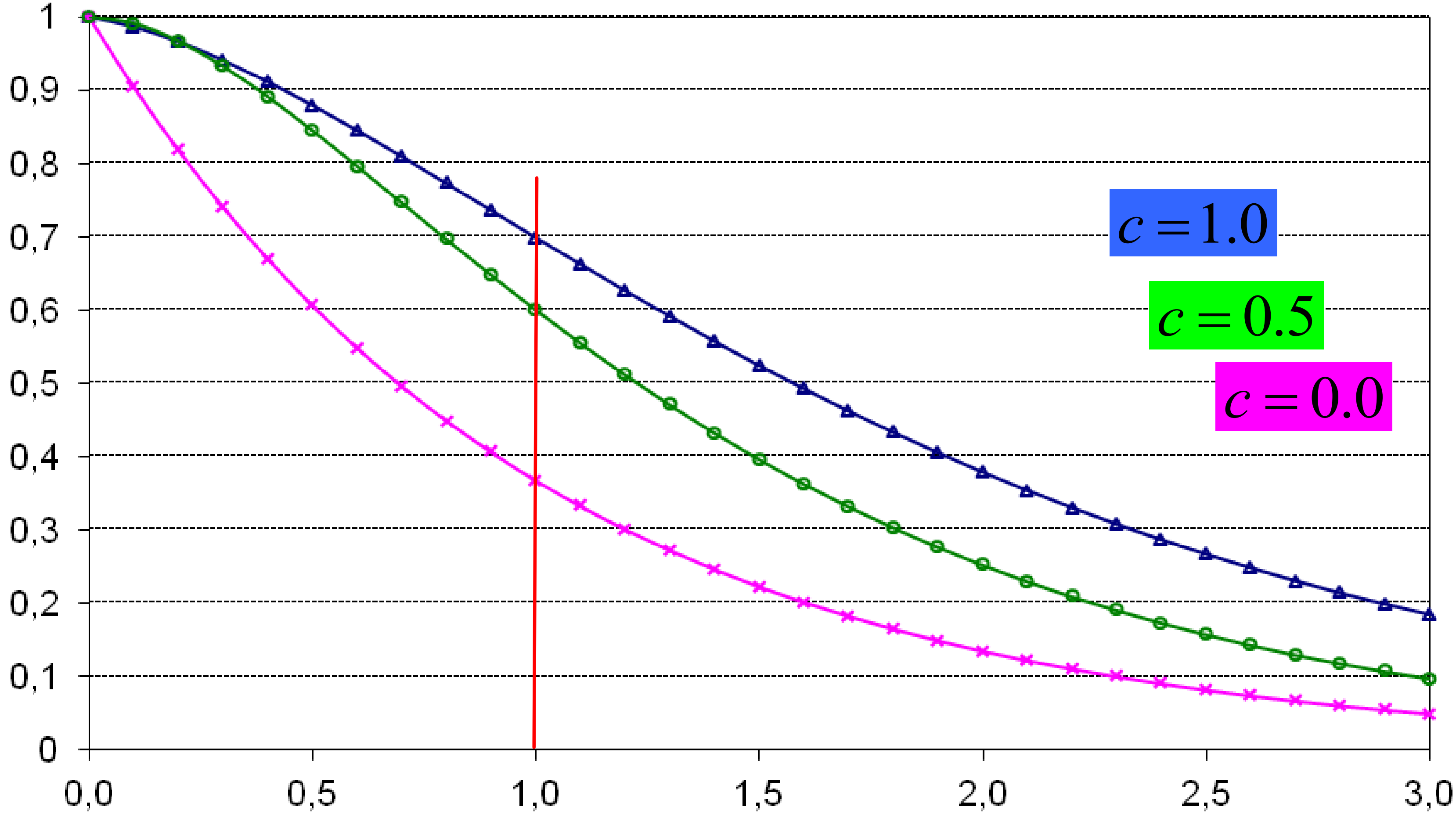
in base al valore di c mi riconduco al caso 0) oppure 1)

se c=0 ho che $R(t) = e^{-\lambda t}$

Effetto del fattore “c” sulla curva di affidabilità

Si nota dal grafico che c influenza l'affidabilità, se si vuole un'affidabilità del 50%, con c=0 si ha fino a circa 750 ore mentre con c=1, si ha fino a 1500 ore. Praticamente il doppio della vita!

$$R(t) = \left(1 + c\lambda\, t\right) e^{-\lambda t}, \quad t \geq 0$$



Da notare come a 1000 ore l'affidabilità non ha divario molto grande tra c=0.5 e c=1

Illustrazione del “peso” della commutazione (1/2)

$$R_{PAR}(t) = \left(e^{-\lambda \cdot t} + e^{-\lambda \cdot t} - e^{-(\lambda+\lambda) \cdot t}\right)$$

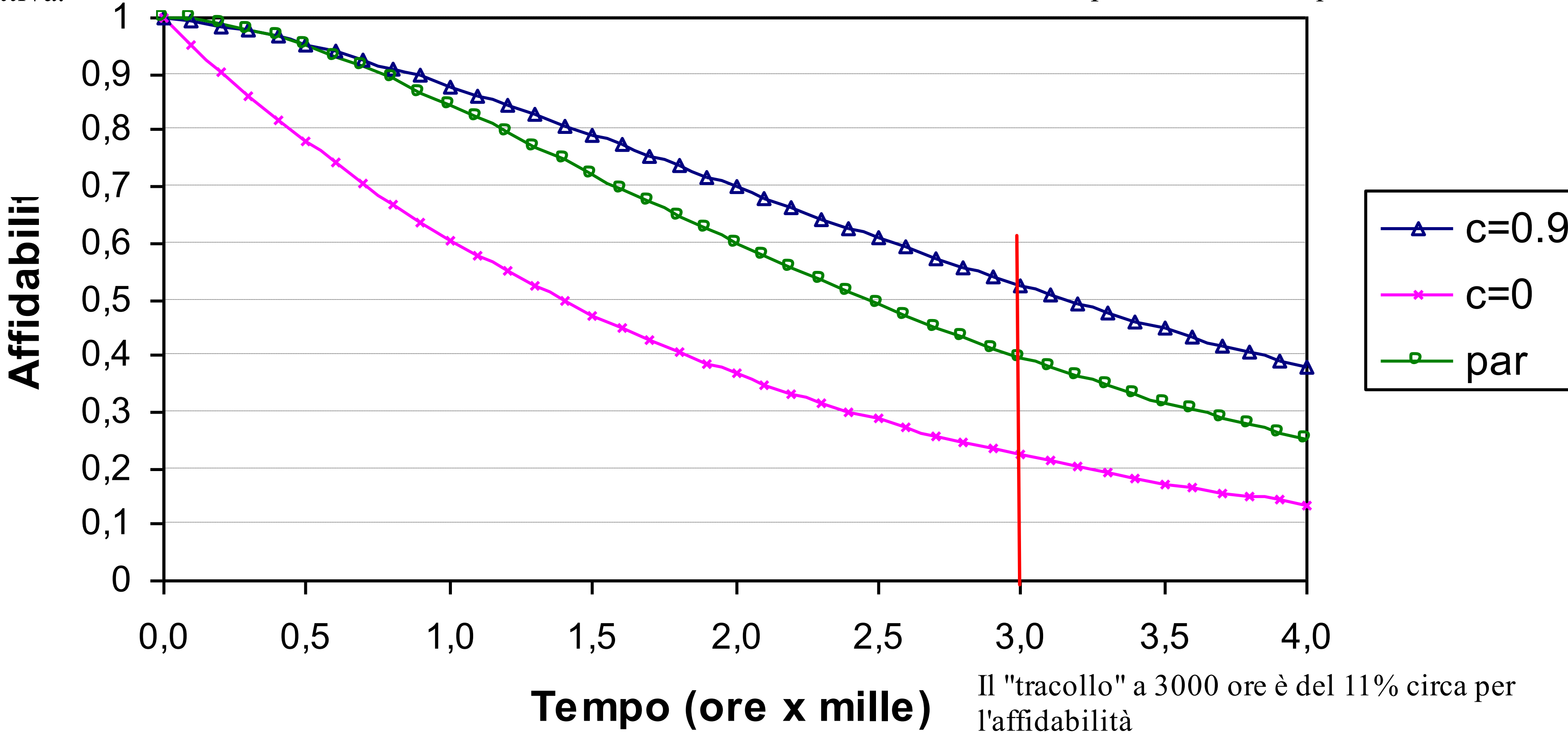
$$R_{RIS}(t) = (1 + c\lambda t)e^{-\lambda t}, \quad t \geq 0$$

I parallelo avrei un sistema m-out-of-n. Con buone prestazioni e cosidetta ridondanza attiva.

Secondo componente in parallelo o di riserva ?
(risultato con c = 0,9 e 1/lambda = 2000 ore)

MEGLIO IN RISERVA!

fai più strada con riserva ma ottieni meno prestazioni sul lungo termine, infatti all'inizio le prestazioni sono quasi identiche



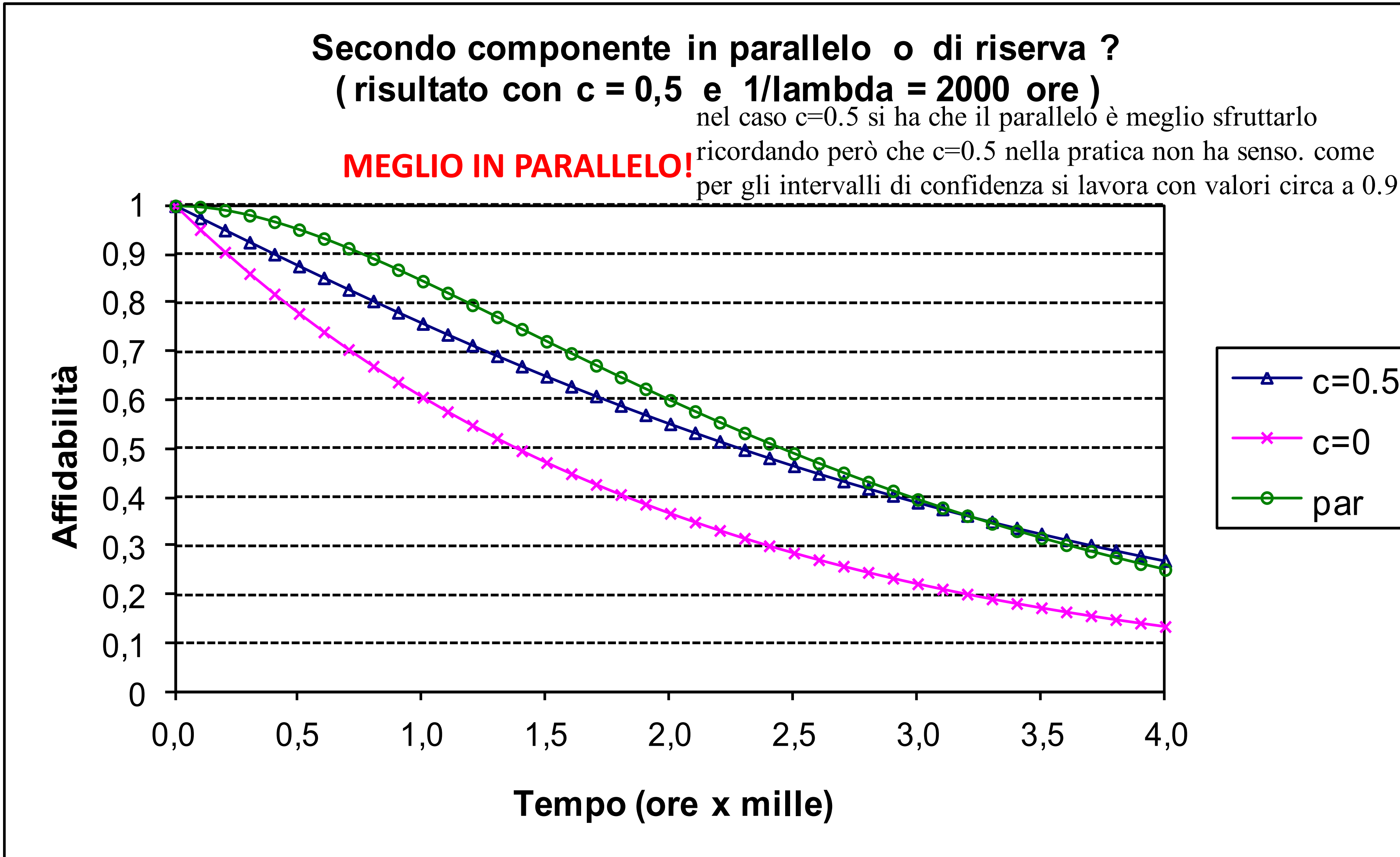
Il "tracollo" a 3000 ore è del 11% circa per l'affidabilità

Illustrazione del "peso" della commutazione (2/2)

$$R_{PAR}(t) = \left(e^{-\lambda \cdot t} + e^{-\lambda \cdot t} - e^{-(\lambda+\lambda)\cdot t}\right)$$

$$R_{RIS}(t) = (1 + c\lambda t)e^{-\lambda t}, \quad t \geq 0$$

nel caso c=0.5 si ha che il parallelo è meglio sfruttarlo ricordando però che c=0.5 nella pratica non ha senso. come per gli intervalli di confidenza si lavora con valori circa a 0.9



Esercizio: MTTF del sistema "1 + 1 di riserva"

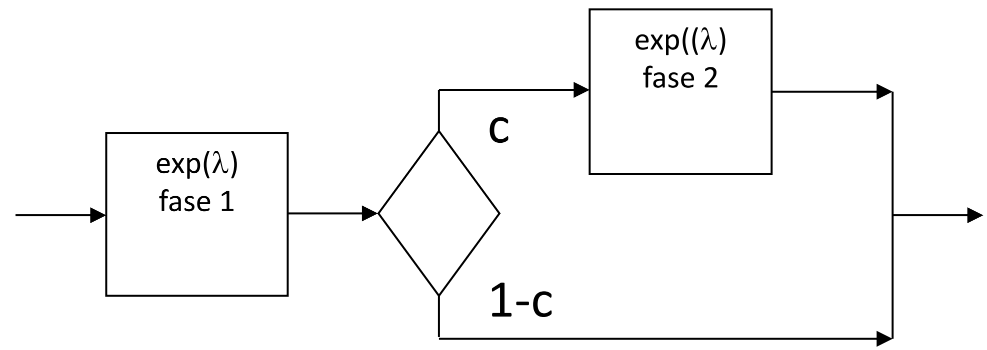
$$MTTF = \int_{t=0}^{\infty} R(t)\,dt = \int_{t=0}^{\infty} (1 + c\lambda t)e^{-\lambda t} \cdot dt =$$

$$= \ldots = \frac{1}{\lambda}(1-c) + \left(\frac{1}{\lambda} + \frac{1}{\lambda}\right) \cdot c$$

tempo al guasto componente attivo o tempo di vita del sistema se la commutazione non riesce

tempo al guasto componente attivo più tempo al guasto della riserva con commutazione perfetta

Il risultato si può ottenere anche applicando la formula del valore atteso totale al seguente diagramma a blocchi che illustra la durata della vita del sistema:



Infatti, (1/λ) è il valore atteso condizionato dall'evento "la commutazione non riesce" e (1-c) è la probabilità dello evento condizionante; mentre (2/λ) è il val atteso condizionato dall'evento "la commutazione riesce" e (c) è la probabilità che la commutazione riesca.

Esercizio: MTTF del sistema "2 in parallelo"

ricordando dalle statistiche dell'ordinamento che la distribuzione del massimo di due durate è pari al prodotto delle distribuzioni

$$X_{PAR} \hat{=} \max\{X_1, X_2\} \quad \Rightarrow \quad F_{X\,PAR}(t) = F_{X_1}(t) \cdot F_{X_2}(t)$$

$$F_{X\,PAR}(t) = \left(1 - e^{-\lambda \cdot t}\right)\left(1 - e^{-\lambda \cdot t}\right)$$

$$= 1 - \left(e^{-\lambda \cdot t} + e^{-\lambda \cdot t} - e^{-(\lambda+\lambda)\cdot t}\right)$$ ottengo la mortalità

$$E[X_{PAR}] = \int_{t=0}^{\infty} \Big(\underbrace{1 - F_{X_{PAR}}(t)}_{\text{affidabilità}}\Big) dt$$

valore atteso del 1° componente

valore atteso del 2° componente

valore atteso del tempo al primo guasto, ma non so di chi

$$= \int_{t=0}^{\infty} \left(e^{-\lambda t} + e^{-\lambda t} - e^{-(\lambda t + \lambda t)}\right) dt \quad = \frac{1}{\lambda} + \frac{1}{\lambda} - \frac{1}{\lambda + \lambda}$$

Esercizio: Confronto fra i due MTTF

In definitiva, fissando $1/\lambda=2000h$ e $c=0.9$:

$$MTTF_{PAR} = \frac{1}{\lambda} + \frac{1}{\lambda} - \frac{1}{2\lambda} = (1+0.5)\cdot\frac{1}{\lambda} = 3000\, h$$

$$MTTF_{RIS} = \frac{1}{\lambda}(1-c) + \frac{2}{\lambda}\, c = (1+c)\cdot\frac{1}{\lambda} = 3800\, h$$

800 ore in più
circa 27%

Dunque:

finché risulta $c > 0.5$ =>> $MTTF_{RIS} > MTTF_{PAR}$

Esempio: faccio un giro con il motoscafo con 2 serbatoi. Basandomi sul consumo ad ore se vado piano passo più tempo in acqua altrimenti se accelero e vado a forti velocità finisco prima il carburante e passo meno tempo sul motoscafo. Esempio informatico: ho da concludere un job, ed ho 2 processori a disposizione la cui velocità si può modulare, allora potrei o fare meno calcoli al secondo ma per più tempo oppure potrei spremere tutta la loro potenza di calcolo per calcolare più in fretta ma in un tempo inferiore. Ad esempio ho 2 GPU e voglio minare un bitcoin in 30 minuti, invece che in 4 ore perché poi arriva la bolletta della luce cara

Problema centrale nel corso

IL PROBLEMA DEL RITARDO A LEZIONE

problema di sincronia

Il professore di "Processi Aleatori e Ingegneria del Teletraffico" arriva a lezione il venerdì mattina alla prima ora con un ritardo aleatorio che va da 0 a 20 minuti massimo, ma arriva!

Lo studente di Arcavacata arriva invece con un ritardo medio di 5 minuti e distribuito secondo il modello esponenziale.

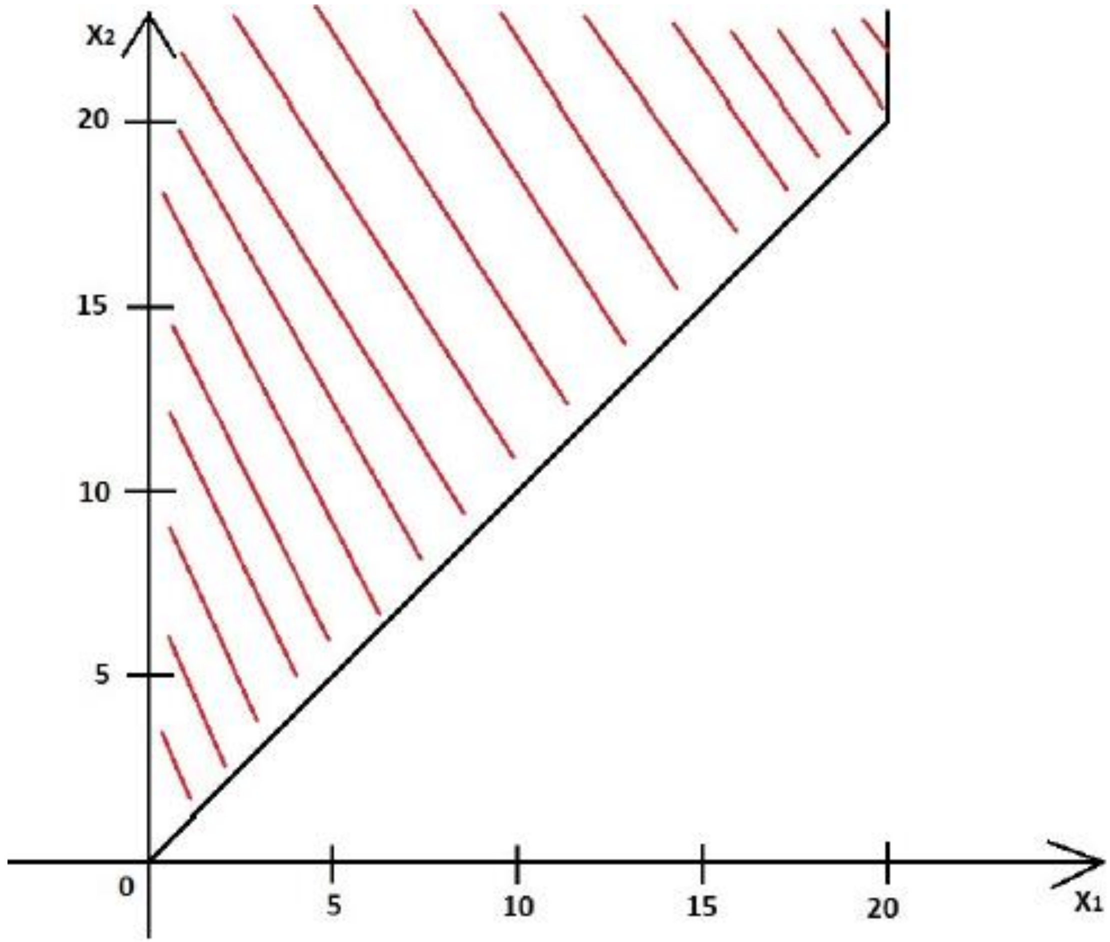
dunque non posso usare il modello di Weibull
uso l'esponenziale con lambda pari 5

Si chiede di:

1) Individuare l'area (A) che corrisponde all'evento di interesse così definito: "lo studente arriva dopo il professore"; X1 sarà il docente X2 sarà lo studente. Ci interessa X2<X1
2) Adottando una funzione di densità lineare, ma compresa fra 0 e 20 minuti per il ritardo del professore, calcolare la probabilità dell'evento di interesse.

QUESITO 1:

Considerando che il docente (X_1) arriva con un ritardo espresso con una densità lineare e che lo studente (X_2) arriva con un ritardo distribuito esponenzialmente con, in media, 5 minuti di ritardo (ma potrebbe anche non arrivare!), si vede dalla seguente figura qual è l'area che corrisponde all'evento desiderato:



QUESITO 2:

I MODELLI DEL RITARDO PER DOCENTE E STUDENTE

Densità della variabile aleatoria X_1 "ritardo del docente":

$$\boxed{f_{X_1}(t) = k \cdot t, \quad \text{con la costante k dalla seguente "docente arriva sicuramente entro 20 minuti"}}$$

$$\boxed{\int_{t=0}^{t=20} k \cdot t \cdot dt = 1} \longrightarrow k \cdot \left[\frac{t^2}{2} \right] \Bigg|_{t=0}^{t=20} = 1 \quad \Rightarrow k = \frac{1}{200}$$

Densità della variabile aleatoria X_2 "ritardo dello studente":

$$E[X_2] = 5 \quad \Rightarrow f_{X_2}(t) = \frac{1}{5} \cdot e^{-\frac{t}{5}} \qquad \text{"inversa" esponenziale}$$

QUESITO 2:

PROBABILITA' CHE LO STUDENTE ARRIVI DOPO IL DOCENTE

risolvo applicando la distribuzione totale

$$P(X_2 > X_1) =$$

condizionata — condizionante

$$= \int_{t=0}^{\infty} \lim_{dt \to 0} P\{X_2 > t \mid t < X_1 \le t + dt\} \cdot P\{t < X_1 \le t + dt\}$$

$$= \int_{t=0}^{\infty} \left[1 - F_{X_2|X_1}(t)\right] \cdot f_{X_1}^{(m)}(t)dt$$

ma professore e studente sono correlati o indipendenti?

$$\text{ipotesi di indipendenza} \Rightarrow \quad = \int_{t=0}^{\infty} \left[1 - F_{X_2}(t)\right] \cdot f_{X_1}(t)dt$$

$$\boxed{= \int_{t=0}^{t=20} e^{-\frac{t}{5}} \cdot \frac{t}{200} dt = \frac{1}{40} \cdot \int_{t=0}^{t=20} \frac{t}{5} e^{-\frac{t}{5}} \cdot dt}$$

CALCOLO:

$$P(X_2 > X_1) = \frac{1}{40} \cdot \int_{t=0}^{t=20} \frac{t}{5} e^{-\frac{t}{5}} \cdot dt$$

Ponendo $x \triangleq t/5 \Rightarrow dt = 5dx \;\Rightarrow\; \dfrac{5}{40} \displaystyle\int_{x=0}^{x=4} xe^{-x} dx$

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

$$\boxed{\begin{aligned} P(X_2 > X_1) &= \frac{1}{8} \cdot \left\{ \left[-xe^{-x} \right]_0^4 + \left[-e^{-x} \right]_0^4 \right\} = \\ &= \frac{1}{8} \cdot \left\{ -4e^{-4} - e^{-4} + 1 \right\} \\ &= \frac{1}{8} \cdot \left\{ -0,0732 - 0,0183 + 0,9084 \right\} = 0,113 \end{aligned}}$$

E se il docente ritarda al max 10minuti?

t=10 e k=1/100 insomma la metà

QUESITO 2: APPROCCIO ALTERNATIVO, integrando la congiunta

più immediato

VALE PERCHÈ IPOTESI INDIPENDENZA

$$P(X_2 > X_1) = \int_A f_{X_1X_2}(t_1,t_2)\cdot dt_1 dt_2$$

$$A \triangleq \{0 \le t_1 \le 20 \cap t_1 \le t_2 < \infty\}$$

$$= \int_{t_1=0}^{20} \int_{t_2=t_1}^{\infty} f_{X_1X_2}(t_1,t_2)\cdot dt_1 dt_2$$

RICORDANDO:

GRAZIE ALL'INDIPENDENZA
(lezione congiunta)

$$f_{X_1,X_2}(t_1,t_2) = f_{X_1}(t_1)\cdot f_{X_2}(t_2)$$

con: $f_{X_1}(t_1) = \dfrac{t_1}{200}$, $f_{X_2}(t_2) = \dfrac{1}{5}e^{-\frac{t_2}{5}}$

SI OTTIENE:

$$P(X_2 > X_1) = \int_{t_1=0}^{20} \int_{t_2=t_1}^{\infty} \frac{t_1}{200}\cdot\frac{e^{-\frac{t_2}{5}}}{5}\cdot dt_1\cdot dt_2$$

CALCOLIAMO L'INTEGRALE INTERNO:

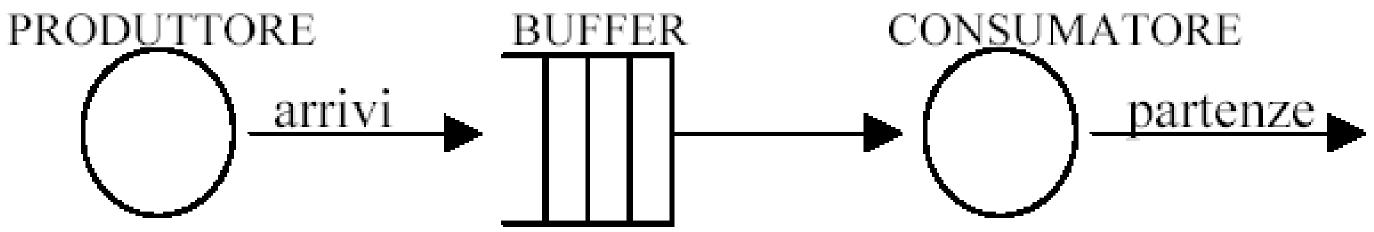
$$\int_{t_2=t_1}^{\infty} \frac{t_1}{200} \cdot \frac{e^{-\frac{t_2}{5}}}{5} \cdot dt_2 = \lim_{t\to\infty} \int_{t_2=t_1}^{t} \frac{t_1}{200} \cdot \frac{e^{-\frac{t_2}{5}}}{5} \cdot dt_2$$

$$= \frac{t_1}{200} \cdot \lim_{t\to\infty} \left[\left(1 - e^{-\frac{t}{5}} \right) - \left(1 - e^{-\frac{t_1}{5}} \right) \right]$$

$$= \frac{t_1}{200} \cdot e^{-\frac{t_1}{5}} = \frac{1}{40} \cdot t_1 \cdot \frac{1}{5} e^{-\frac{t_1}{5}}$$

E RISCRIVIAMO: $$P(X_2 > X_1) = \frac{1}{40} \int_{t_1=0}^{20} \frac{t_1}{5} e^{-\frac{t_1}{5}} \cdot dt_1$$

Il modello produttore-consumatore (P-C)



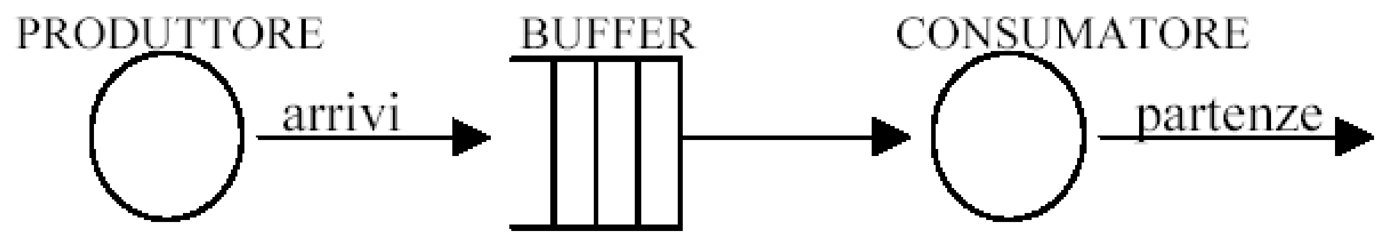
Il “produttore” implementa un **generatore** di “entità” destinate al “consumatore” e temporaneamente depositate nel buffer.

La comunicazione fra i due oggetti avviene secondo un **protocollo** che regola le “attività” di produzione e consumo.

La durata delle attività è **aleatoria** e scandisce il passaggio delle entità attraverso il sistema, determinando un **flusso aleatorio**.

Il problema della distribuzione del tempo di attesa

Ipotesi: il server non viene interrotto da nessun altro lavoro e non rifiuta richieste



Ipotesi: clienti infiniti e che sono caratterizzati solo dal ritmo di arrivi lambda

Non vado a studiare il comportamento di W(m), ad esempio W(10) ovvero del decimo consumatore. Ma eseguo uno studio a lungo termine

Var. al. sul lungo termine:

quindi ne studio una sola

$$W \hat{=} \text{"}\textit{attesa nel buffer}\text{"}$$

$$F_W^{(m)}(t) \hat{=} \Pr[W \leq t]$$

$$= \sum_{n=0}^{\infty} \Pr[W \leq t \cap N = n]$$

è una congiunta. Attendo meno di t e ci siano n utenti presenti

$$= \sum_{n=0}^{\infty} \Pr[W \leq t \mid N = n] \cdot \Pr[N = n]$$

l'attesa W è una marginale condizionata dal numero di utenti presenti N

$$\boxed{= \sum_{n=0}^{\infty} F_{W|N}(t \mid n) \cdot \Pr[N = n]}$$

distribuzione totale

N$\hat{=}$ "utenti presenti" sul lungo termine, in corrispondenza di istanti completamente casuali

trovo n clienti in coda, qual è la probabilità che vengo servito entro un tempo?

Se arrivi e partenze sono esponenziali $F_{W|N}(t|n)$ è modello di Erlang

C'è differenza tra utenti trovati ed utenti presenti. Il numero di utenti trovati dipendono dal profilo probabilistico che si immagina di avere, estraendo il mio istante di arrivo. Quindi a t1 trovo n utenti mentre a t2 trovo m utenti. Gli utenti presenti non dipendono dall'arrivo ma da un osservatore casuale. Se considero uno stimatore casuale gli istanti di arrivo sarebbero distribuiti in maniera equiprobabile. Nel problema della distribuzione del tempo di attesa considero gli arrivi in maniera completamente casuale, (Poissoniani: distribuiti uniformemente nel orizzonte temporale) perciò il punto di vista dell'osservatore casuale coincide con il punto di vista del singolo utente, ovvero statisticamente indistinguibili. Se non fosse dovrei impiegare una simulazione monte carlo

Il teorema (sulla distribuzione del tempo di attesa)

μ velocità di servizio ed anche produttività limite, quando il server è pieno

IPOTESI: λ ritmo degli arrivi

$\lambda = \rho\mu$ equazione di continuità, ciò che entra esce o meglio disciplina first in first served (F.I.F.S)

Il modello P-C ha un solo server gestito con disciplina "F.I.F.O.",

"1/λ" >0 è il valore atteso del tempo tra due arrivi consecutivi, A,

"1/μ" >0 è la durata attesa del servizio reso al generico cliente, S,

probabilità che arrivi un nuovo cliente entro t

$$\left\{ A : F_A(t) = 1 - e^{\lambda t}, t \geq 0 \right\}$$

probabilità che il servizio abbia durata meno di t

$$\left\{ S : F_S(t) = 1 - e^{\mu t}, t \geq 0 \right\}$$

Altra IPOTESI

TESI: è $\boxed{\text{valida ses:} 1 > (\lambda / \mu) \triangleq \rho}$

ρ è sia numero medio di arrivi (valore atteso) durante un servizio sia il fattore di utilizzo della risorsa ma non è produttività! Se il fattore è molto alto si ha molta richiesta

$$\boxed{\left(\begin{aligned} \rho = \Pr\{\text{"}\underset{\text{almeno uno}}{\text{utenti} \ \geq \ 1}\text{"}\} \\ = \Pr\{\text{"attesa} > 0\text{"}\} \end{aligned} \right)}$$

$$1. \quad P_N(n) = (1-\rho) \cdot \rho^n, \quad n \geq 0$$

ρ è prob che ci sia almeno 1 è uguale a quella che il server stia lavorando

$$2. \quad F_W(t) = F_{W|N=0}(t) P_{N=0} + \sum_{n=1}^{\infty} F_{W|N=n}(t) P_N(n)$$

$$= 1 \cdot (1-\rho) + (1 - e^{-\mu(1-\rho)t}) \cdot \rho = 1 - \rho e^{-\mu(1-\rho)t}$$

Si dimostra che si ottiene il modello geometrico

$$P_N(n) = (1-\rho)\cdot\rho^n \quad n \geq 0$$ se n=0 non c'è nessuno

probabilità server è libero/buffer vuoto

probabilità di occupazione del buffer con n clienti

E che il modello probabilistico è il seguente

Probabilità evento condizionante che non ci sia nessuno

attesa condizionata di chi non aspetta

probalità che non ci sia nessuno

attesa condizionante di chi aspetta

$$F_W(t) = F_{W|N=0}(t)P_{N=0} + \sum_{n=1}^{\infty} F_{W|N=n}(t)P_N(n) = 1\cdot(1-\rho) + (1-e^{-\mu(1-\rho)t})\cdot\rho = 1-\rho e^{-\mu(1-\rho)t}$$

Distribuzione dell'attesa (totale)

Distribuzione condizionata che non ci sia nessuno

Erlang

se non c'è nessuno l'unica realizzazione è pari a 0. dunque tutta la prob. è concentrata

attesa in generale che NON è una funzione esponenziale

Il modello P_C e l'attesa dal medico di base

(vale SEMPRE l'ipotesi di arrivi casuali)

Il dott. Tizio, medico di base, valuta che il flusso degli arrivi di pazienti al suo studio possa essere quantificato in 3 pazienti all'ora, in media su tutto l'intervallo giornaliero
lambda
di apertura dell'ambulatorio. L'ambulatorio si compone di una sala d'attesa e dello studio per le visite.

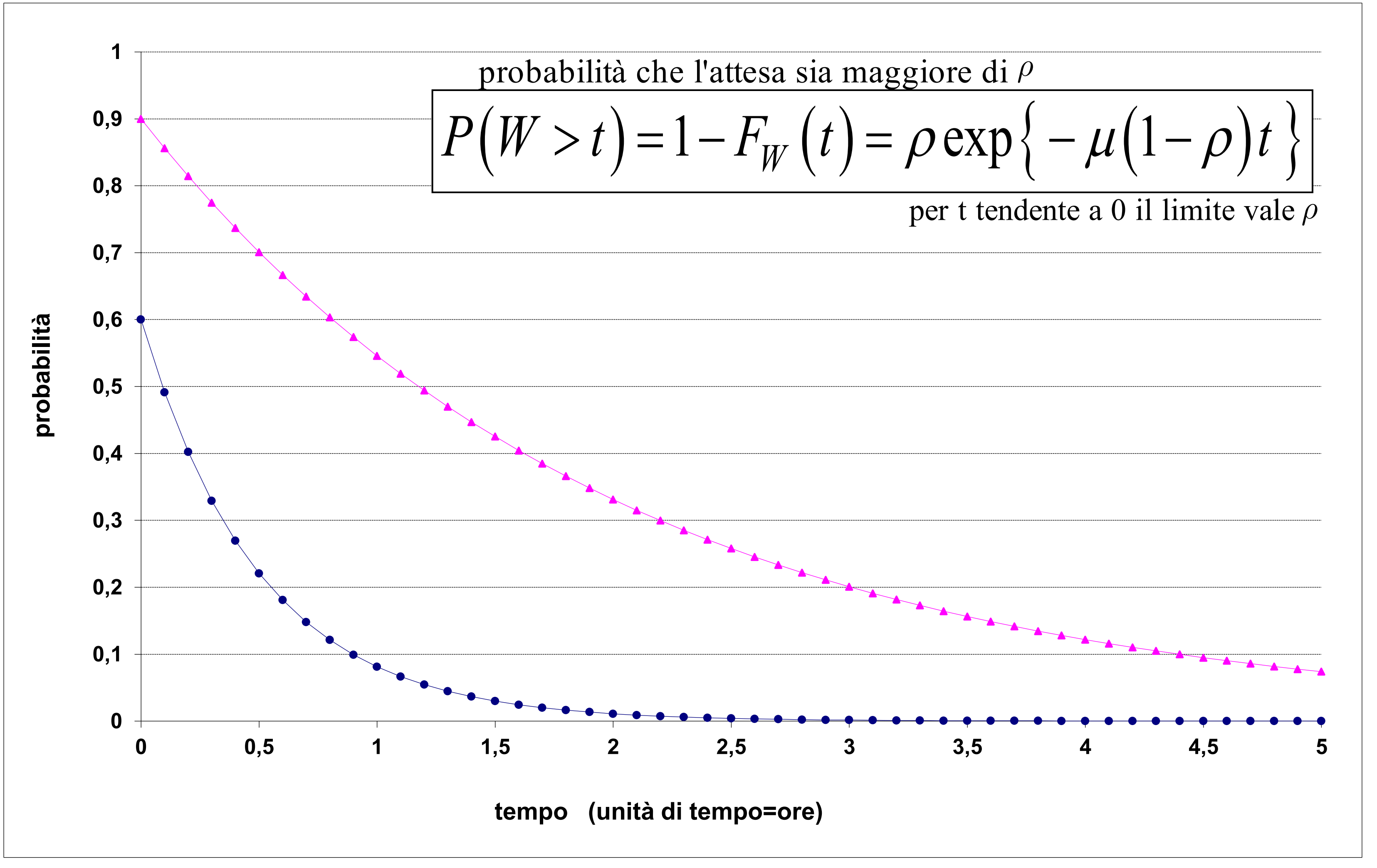
mu=12*5=60

I pazienti vengono visitati in ordine di arrivo e ogni visita dura mediamente 12 minuti

Verificare che è praticamente nulla la probabilità che un paziente debba aspettare più di 2 ore prima di essere visitato e che, d'altra parte, il 78% dei pazienti in arrivo aspetta non più di mezz'ora.

Considerato che il dott. Tizio non andrà in ferie d'estate, valutare se può accettare nel proprio ambulatorio anche parte dei pazienti del dott. Caio che, invece, andrà in ferie. Valutando che si avrebbe un incremento del ritmo orario medio degli arrivi da $3h^{-1}$ a
incremento 50%
$4{,}5h^{-1}$ nell'ambulatorio del dott. Tizio, verificare che la probabilità di aspettare più di 2 ore prima di essere visitati aumenterebbe da quasi zero al valore 0,33.

Probabilità di attesa dal medico di base



la curva blu rappresenta un server più efficente

DURATA DEI SERVIZI: $\left\{ S_i = S_j = S \hat{=} \exp(\mu), \ \ \forall j = 1,2,\ldots,n \right\}$

$$FIFO \Rightarrow \ "W \mid n" = \sum_{i=1}^{n} S_i$$

INDIPENDENZA DEI SERVIZI

$$\Pr\left\{ \sum_{i=1}^{n} S_i \leq t \right\} = 1 - e^{-\mu t} \sum_{i=0}^{n-1} \frac{(\mu t)^i}{i!} \ \hat{=} \ F_{W|N}(t \mid n)$$

Modello di Erlang

Ecco ora il procedimento con cui si ricava la distribuzione del tempo d'attesa in coda, $F_W(t)$:

$$F_W(t) = \sum_{n=0}^{\infty} F_{W|n}(t) P_n = F_{W|0}(t) P_0 + \sum_{n=1}^{\infty} F_{W|n}(t) P_n$$

Quando la risorsa-servente è libera, la variabile attesa condizionata avrà la sola realizzazione nulla, dunque:

$$F_{W|n}(t) = 1, \quad n = 0 \quad \text{e pertanto:} \quad F_{W|0}(t) P_0 = 1(1-\rho)$$

Viceversa, quando sono presenti n utenti di fronte al nuovo arrivato, risulterà:

$$F_{W|n}(t) = \ 1 - e^{-\mu t} \sum_{i=0}^{n-1} \frac{(\mu t)^i}{i!}, \quad n \geq 1$$

Allora:

La gara fra due durate esponenziali indipendenti

IPOTESI

(hanno cominciato assieme)

X_1 e X_2 sono le durate di due attività in parallelo svolgimento:

se sono indipendenti

$$F_{X_2|X_1}(t) = F_{X_2}(t)$$

$$P(X_1 < X_2) = \Pr(X_2 > t | t < X_1 \leq t + dt\ )$$

probalità condizionata

distribuzione totale

$$= \int_{t=0}^{\infty} \left[1 - F_{X_2|X_1}(t) \right] \cdot f_{X_1}^{(m)}(t) dt$$

sopravvivenza

probabilità variabile condizionante che duri tra t e t+dt

$$f_{X_1}^{(m)} = \frac{f_{X_1,X_2}}{f_{X_2/X_1}} = f_{X_1}$$

(INDIPENDENZA STOCASTICA)

caso più semplice

Per due modelli ESPONENZIALI con parametri μ_1 e μ_2:

$$P(X_1 < X_2) = \int_{t=0}^{\infty} e^{-\mu_2 t} \cdot \mu_1 e^{-\mu_1 t} dt$$

$$= \frac{\mu_1}{\mu_1 + \mu_2} \int_{t=0}^{\infty} (\mu_1 + \mu_2) e^{-(\mu_2 + \mu_1)t} dt = \frac{\mu_1}{\mu_1 + \mu_2}$$

come visto in precedenza

velocità del primo modello fratto la somma delle due velocità

La durata della gara e il vincitore

Siano

X_1, X_2 esponenziali e indipendenti

$$\Rightarrow P(X_1 < X_2) = \frac{\mu_1}{\mu_1 + \mu_2}$$

estensione possibile a 4 partecipanti: $X_1 < X_2 \cap X_3 \cap X_4$

Se aggiungiamo che: $Z \triangleq \min(X_1, X_2)$ statistica dell'ordinamento

probabilità che X_1 finisca entro t

$$\Rightarrow F_Z(t) = 1 - \underbrace{(1 - F_{X_1}(t))}_{\text{probabilità che } X_1 \text{ finisca dopo t}} \cdot \underbrace{(1 - F_{X_2}(t))}_{\text{probabilità che } X_2 \text{ finisca dopo t}} = 1 - e^{-(\mu_1+\mu_2)\mathrm{t}}$$

$$P(Z \leq t \cap X_1 < X_2) =$$

$$= \underbrace{P(Z \leq t \mid X_1 < X_2)}_{\text{condizionata}} \cdot \underbrace{P(X_1 < X_2)}_{\text{condizionante}}$$

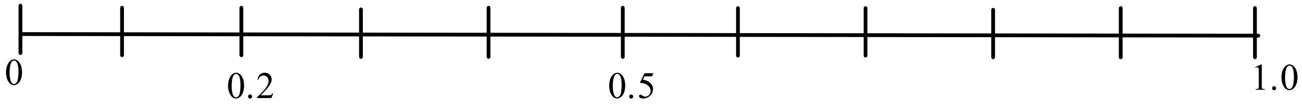
? qui si gioca l'indipendenza

$$= P(Z \leq t) \cdot P(X_1 < X_2)$$

$$= (1 - e^{-(\mu_1+\mu_2)t}) \cdot \frac{\mu_1}{\mu_1 + \mu_2}$$

Ecco la probabilità che il primo fenomeno finisca entro t e che sia proprio il fenomeno 1

Esperimento di Bernoulli generalizzato per il caso generico di n exp in gara
Genero un numero casuale ed in base al suo valore vedo in quale intervallo cade. Esempio:



Se casca in 0.2 allora ho che la prima attività a finire è $\frac{\mu_1 = 0.2}{\sum_{j=1}^{n} \mu_j}$

Se casca in 0.3 il primo a finire è $\frac{\mu_2 = 0.3}{\sum_{j=1}^{n} \mu_j}$

La distribuzione iperesponenziale

non è una convoluzione

$$\boxed{Y \hat{=} \sum_{i=1}^{n} \alpha_i \cdot Y_i \;;}$$

ma combinazione convessa di exp

$$\begin{cases} \alpha_i > 0, \quad \sum_{i=1}^{n} \alpha_i = 1; \\ F_{Y_i}(y) \hat{=} 1 - \exp\{-\lambda_i y\} \end{cases}$$

se exp tutte uguali non si cambia niente

COMBINAZIONE CONVESSA DI EXPONENZIALI

Interpretazione:

$$P_X(i) \hat{=} \alpha_i, \quad i = 1, \ldots, n$$

prob. eventi condizionanti

Distribuzione totale:

$$F_Y(y) = \sum_{i=1}^{n} \underbrace{F_{Y|X=i}(y \mid i)}_{\text{condizionata della durata}} \cdot \underbrace{P_X(i)}_{\text{condizionante}}$$

Risultato:

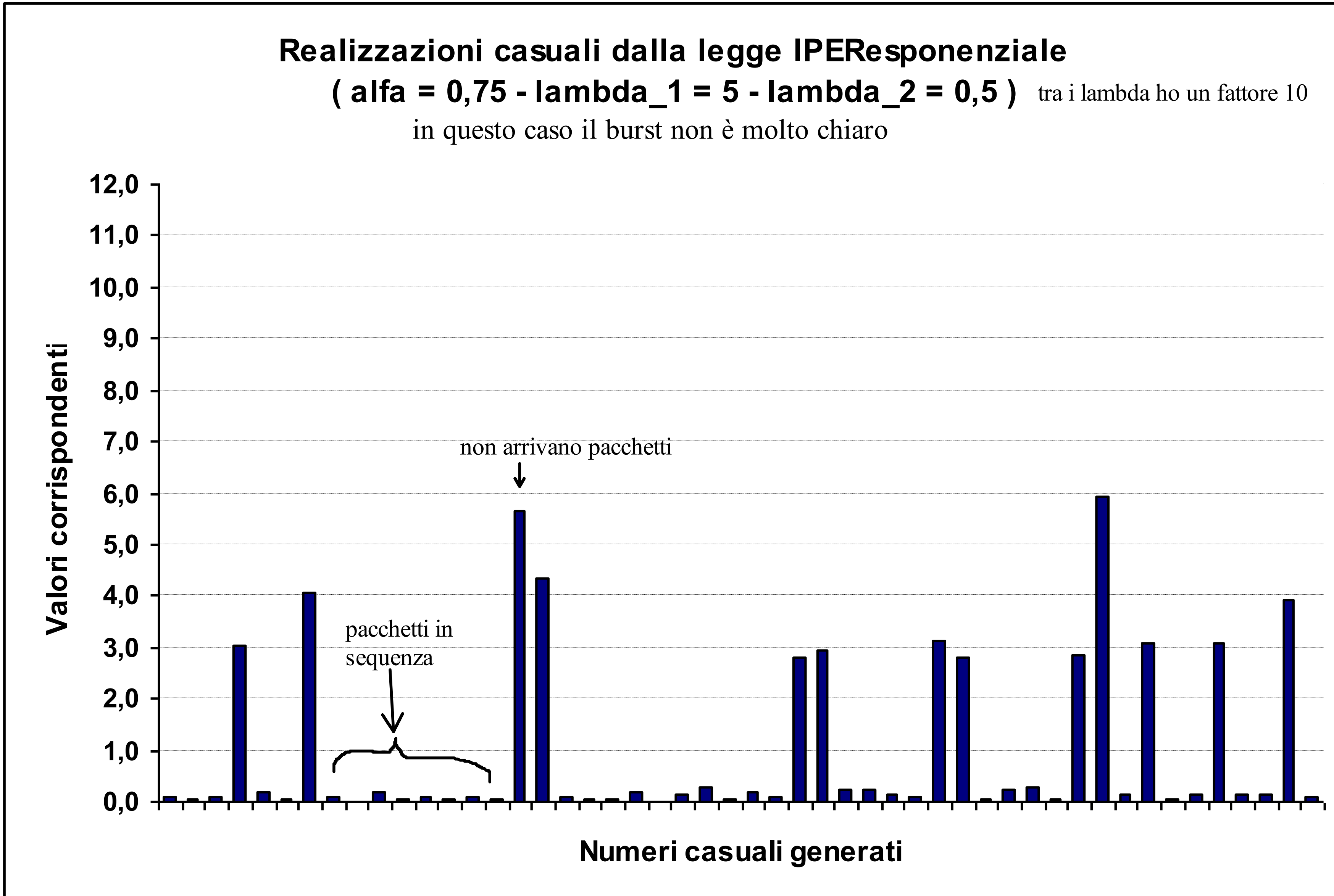
$$F_Y(y) = \sum_{i=1}^{n} \alpha_i \left(1 - \exp\{-\lambda_i y\}\right)$$

Dunque l'iperesponenziale nasce dalla distribuzione totale

Quindi la densità:

$$f_Y(y) = \sum_{i=1}^{n} \alpha_i \lambda_i \exp\{-\lambda_i y\}$$

Ricordando che l'iperexp mostra l'effetto di burst nelle reti oppure di esecuzione di processi in un OS

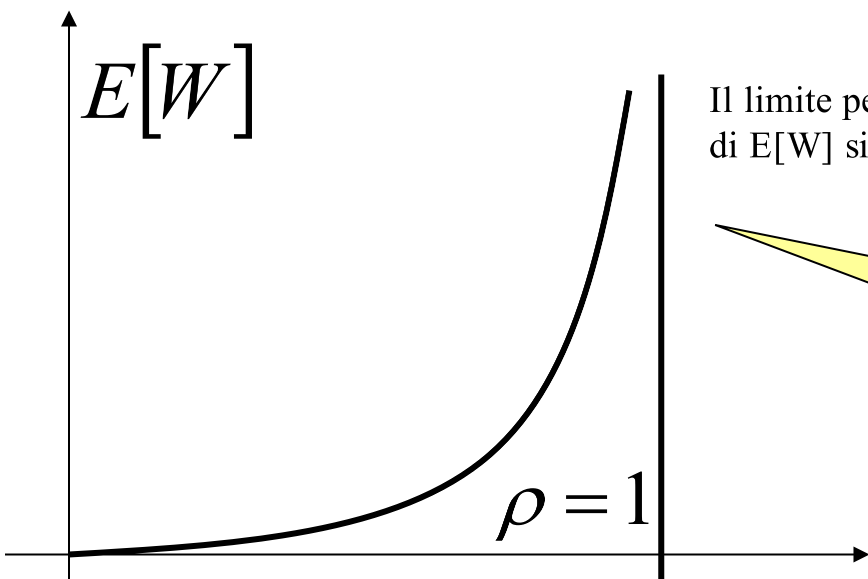


… ***attesa media:*** $$E[W]=\int_{t=0}^{\infty}\left(1-F_W(t)\right)dt$$

$$=\int_{t=0}^{\infty}\rho e^{-\mu(1-\rho)t}\,dt = \frac{\rho}{\mu(1-\rho)}$$

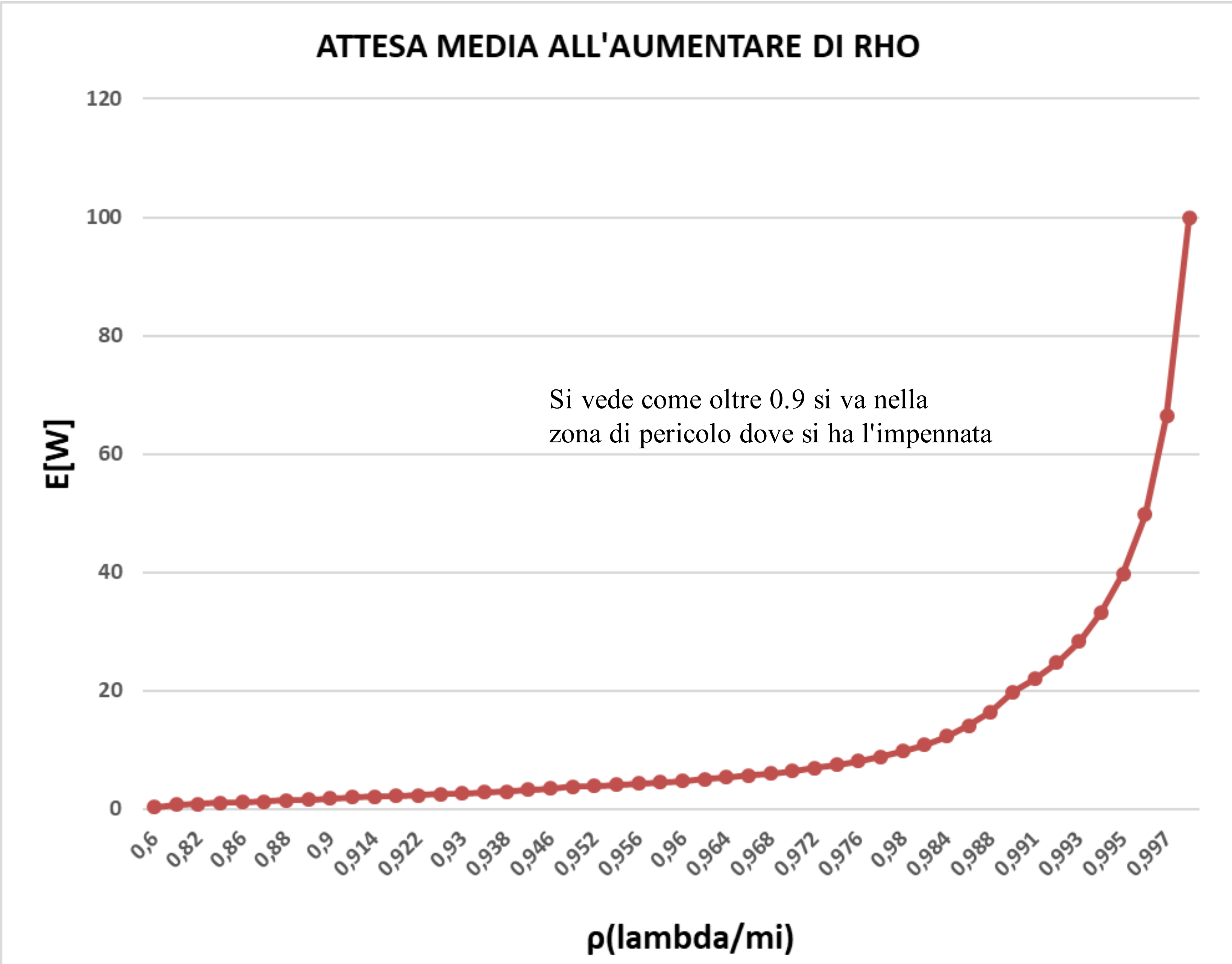
$$E[D] = E[W]+\mu^{-1} = \frac{1}{\mu(1-\rho)}$$

Tempo (Delay) speso nel modello P-C



Il limite per rho tendente a 1
di E[W] si ha un asintoto quindi si attende infinitamente

L'utente trova una coda infinitamente lunga


ATTESA MEDIA ALL'AUMENTARE DI RHO
120
100
80
60
40
20
0
E[W]
Si vede come oltre 0.9 si va nella
zona di pericolo dove si ha l'impennata
0,6
0,82
0,86
0,88
0,9
0,914
0,922
0,93
0,938
0,946
0,952
0,956
0,96
0,964
0,968
0,972
0,976
0,98
0,984
0,988
0,991
0,993
0,995
0,997
ρ(lambda/mi)

Il modello P_C e l’attesa dal medico di base

Il risultato appena verificato non è accettabile per la qualità del servizio che il dott. Tizio vuole garantire a tutti i pazienti, alcuni dei quali sarebbero pure tentati di abbandonare la coda di attesa in ambulatorio. Allora, egli decide di adottare una politica di “rigetto senza ritorno” nei confronti di pazienti in arrivo.

La politica consiste nel chiudere la porta dell’ambulatorio quando si raggiunga il numero di 3 pazienti in attesa oltre quello sotto visita e dirottare presso la guardia medica o un qualsivoglia altro collega chiunque trovi la porta chiusa quando arriva in ambulatorio.

Il modello P_C e l'attesa dal medico di base

Adottando il modello geometrico troncato per la variabile aleatoria "numero di pazienti presenti in ambulatorio" ricalcolare la probabilità che in presenza della condivisione di pazienti (e dunque ritmo medio orario degli arrivi pari a 4,5) un paziente in arrivo non aspetti più di mezz'ora, posto che venga accettato.

Domanda finale riferita al modello a buffer illimitato:

E se volessimo calcolare la probabilità che il tempo totale di permanenza dal dottore (attesa + visita) risulti non superiore alle 2 ore, quando il ritmo degli arrivi sale a 4,5h-1?

Risposta: avremmo bisogno della funzione di distribuzione riferita al tempo totale (attesa + visita), che si dimostra essere la seguente:

$$P\left(W+V\leq t\right)\hat{=}F_{W+V}\left(t\right)=1-exp\left\{-\mu\left(1-\rho\right)t\right\}$$

La troncata a partire dalla geometrica NON troncata

Dopo un tot di clienti si ha un rigetto da parte del serve. È un richiamo sulla geometrica

$$P_n = (1-\rho)\rho^n, \quad n = 0, 1, \ldots, k, k+1, \ldots$$

Evento di interesse:

$$\text{"}esattamente \text{ n } utenti\ presenti \mid al\ più\ k\ possono\ essere\ accettati\text{"} \;\hat{=}\; n \mid n \le k$$

Probabilità evento condizionante:

$$P_{n\le k} = \sum_{n=0}^{k} (1-\rho)\rho^n$$

Probabilità evento d'interesse:

$$P_{n|n\le k} = \frac{P_n}{P_{n\le k}} = \frac{(1-\rho)\rho^n}{\sum_{n=0}^{k}(1-\rho)\rho^n} \qquad n = 0, 1, \ldots, k$$

Derivazione formula finale per $P_{n|n\le k}$ $n = 0,1,\dots,k$

Calcoli per capire non per esame

Si sviluppa: $\boxed{\sum_{n=0}^{k} \rho^n}$ $\boxed{= \sum_{i=0}^{\infty} \rho^n - \sum_{n=k+1}^{\infty} \rho^n}$

$$con \;\; \rho < 1 \;\; e \;\; j \hat{=} n-(k+1) \Rightarrow \qquad = \frac{1}{1-\rho} - \sum_{j=0}^{\infty} \rho^{k+1}\rho^{j}$$

$$= \frac{1}{1-\rho} - \frac{\rho^{k+1}}{1-\rho} = \frac{1-\rho^{k+1}}{1-\rho}$$

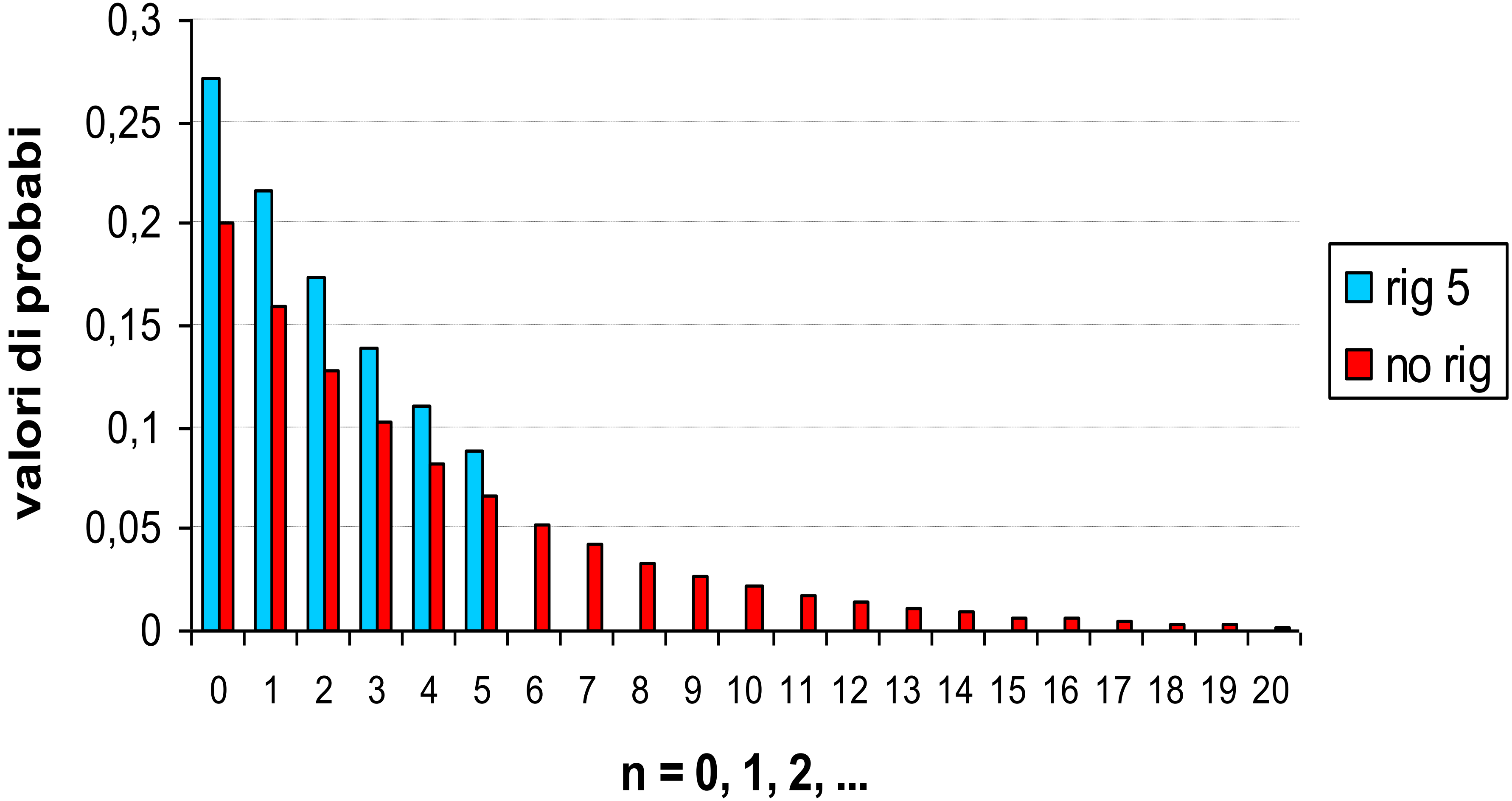
Formula finale:

$$\boxed{P_{n|n\le k} = \frac{\rho^n}{\sum_{n=0}^{k} \rho^n} = \frac{(1-\rho)}{1-\rho^{k+1}} \rho^n}$$

$$n = 0,1,\dots,k$$

Prob stazionaria (utenti presenti = n), con $\rho = 0.8$

Di più sul foglio EXCEL



Formula del valore atteso totale (per var. al. continue)

Si parte dalla definizione di valore atteso della Y:

$$E[Y] \hat{=} \int_{y=0}^{\infty} y \cdot f_Y^{(m)}(y) \cdot dy$$ = passaggio con la congiunta

Estensione del concetto per variabile condizionata

e si esprime la densità marginale della Y come prodotto della densità condizionata di Y|X per la densità marginale della condizionante, X:

$$= \int_{y=0}^{\infty} y \cdot \left[\int_{x=0}^{\infty} f(y \mid x) f_X^{(m)}(x) dx \right] dy \quad = \int_{x=0}^{\infty} \left[\int_{y=0}^{\infty} y \cdot f(y \mid x) dy \right] f_X^{(m)}(x) dx$$

sommo tutti i casi possibili della variabile condizionante

A questo punto, si definisce il valore atteso condizionato:

$$\boxed{E[Y \mid X = x] \quad \int_{y=0}^{\infty} y \cdot f(y \mid x) \cdot dy} < +\infty$$

fisso il valore condizionante a x ad esempio impiego 10 secondi per un processo al variare di x (valore condizionante) si ha una funzione, una curva detta curva di regressione

e si ottiene la seguente:

$$E[Y] = \int_{x=0}^{\infty} E[Y \mid X = x] \cdot f_X^{(m)}(x) \cdot dx$$

valore atteso condizionato per la condizionante sommato su tutti i casi possibili quindi ottengo il valore atteso totale

Nota come formula del valore atteso totale

Formula del valore atteso totale

nel modello produttore consumatore

$$S \hat{=} \text{" } tempo\ di\ consumo \text{"}$$ (non si sa la durata)

o meglio servizio

$$N_S \hat{=} \text{" }\#\ di\ produzioni\ durante\ un\ consumo \text{"}$$

arrivi   servizio

Qual è il numero atteso di arrivi durante un servizio?

Val. att. condizionato:

della durata del servizio

$$\boxed{E[N_S \mid t] = \lambda \cdot t}$$

significa S=t

pongo un vincolo
con un valore fisso

nel protocollo Aloha si ha che
$G = \lambda T$

Val. att. totale

$t < S \leq t + dt = f_S(t)dt$

$$E[N_S] \hat{=} \int_{t=0}^{\infty} E[N_S \mid S = t] \cdot \Pr\{S \in [t, t+dt]\}$$

valore atteso condizionato

$$= \int_{t=0}^{\infty} \lambda t \cdot \mu e^{-\mu t}\, dt \quad = \quad \frac{\lambda}{\mu} \quad \hat{=}\, \rho$$

Num. medio di arrivi durante un consumo

servizio

Ricapitolando quanto detto nella lezione precedente abbiamo:

λ : ritmo degli arrivi

ρ : numero medio di arrivi durante un servizio e probabilità che ci sia almeno un cliente da servire

μ : capacità di servizio del server e produttività massima del server, nel caso in cui, ci sia sempre almeno un cliente che si sta servendo e che arrivino clienti con un ritmo di arrivo imposto, anche se questo porterebbe ad avere code infinitamente lunghe sul lungo termine

$\rho = \dfrac{\lambda}{\mu}$ per definizione $\qquad$ $\rho \cdot \mu$ il prodotto è il ritmo delle uscite, il cosiddeto throughput del sistema

Un sistema in cui $\dfrac{\lambda}{\mu} > 1$ non è ammissibile!

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:

$$E[X+Y]=E[X]+E[Y] \qquad \text{(proprietà di additività)}$$

$$E[cX+d]=cE[X]+d \qquad (\, c \text{ e } d \text{ costanti arbitrarie}\,)$$

e aggiungendo l'ipotesi di indipendenza

$$E[X\,Y]=E[X]E[Y] \qquad \text{(proprietà di moltiplicatività)}$$

Le precedenti rimangono valide passando da due a "n" variabili.

La **varianza** di una variabile aleatoria X è definita come:

$$VAR[X]\equiv E\left[(X-E[X])^2\right]$$

La precedente espressione risulta pari a:

$$VAR[X]=E\left[X^2-2XE[X]+E[X]^2\right]=...=E[X^2]-E[X]^2$$

Con riferimento ad una coppia di variabili aleatorie continue, X e Y, allo spazio delle realizzazioni congiunte $\{(x,y)\in\Re^2 \mid 0\le x<\infty,\ 0\le y<\infty\}$ 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:

$$E[g(X,Y)]\triangleq\int_{x=0}^{\infty}\int_{y=0}^{\infty} g(x,y)f_{X_1X_2}(x,y)\,dy\,dx \quad <\infty$$

Riportata al caso di una coppia di variabili discrete, N e K, con lo spazio delle realizzazioni congiunte $\{(n,k)\mid n=0,1,2,...\ ,\ k=0,1,2,...\}$ e la generica funzione g(N,K), la definizione corrispondente è:

$$E[g(N,K)]\triangleq\sum_{n=0}^{\infty}\sum_{k=0}^{\infty} g(n,k)P_{N,K}(n,k) \quad <\infty$$

La varianza della somma di due variabili aleatorie risulta:

$$VAR[X+Y] \triangleq E\left[((X+Y)-E[X+Y])^2\right]$$

$$= E\left[((X+Y)-E[X]-E[Y])^2\right]$$

$$= E\left[(X-E[X])^2+(Y-E[Y])^2+2(X-E[X])(Y-E[Y])\right]$$

$$= E\left[(X-E[X])^2\right]+E\left[(Y-E[Y])^2\right]+2E[(X-E[X])(Y-E[Y])]$$

$$= VAR[X]+VAR[Y]+2E[(X-E[X])(Y-E[Y])]$$

Generalizzazione della formula della somma:

$$VAR[X_1+X_2+\ldots+X_n]=$$

$$= E\left[\left((X_1+X_2+\ldots+X_n)-E[X_1+X_2+\ldots+X_n]\right)^2\right]=$$
$$= E\left[\left(X_1+X_2+\ldots+X_n-E[X_1]-E[X_2]-\ldots-E[X_n]\right)^2\right]=$$
$$= E\left[\left((X_1-E[X_1])+(X_2-E[X_2])+\ldots+(X_n-E[X_n])\right)^2\right]=$$

$$= VAR[X_1]+\ldots+VAR[X_n]+$$
$$+2\sum_{i=1}^{n-1}\sum_{j=i+1}^{n}E\left[(X_i-E[X_i])(X_j-E[X_j])\right]$$

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: $\lambda_i \neq \lambda_j$, $i \neq j$:

$$\boxed{\exp(\lambda_1)} \longrightarrow \boxed{\exp(\lambda_2)} \longrightarrow \dots \longrightarrow \boxed{\exp(\lambda_n)}$$

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*:

$$f_X(x) = \sum_{i=1}^{n} a_i\, \lambda_i \exp\{-\lambda_i x\} \qquad 0 \le x < \infty$$

e con distribuzione:

$$F_X(x) = \sum_{i=1}^{n} a_i \left(1 - \exp\{-\lambda_i x\}\right) \qquad 0 \le x < \infty$$

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. $X_i \triangleq$ "tempo al guasto del componente i":

$$F_{X_i}(t) = 1 - e^{-\lambda t}, \quad i = 1,2,3$$

- per il componente di riserva invece:

$$F_{X_r}(t) = \begin{cases} 1 - e^{-\lambda t} & attiva \\ 1 - e^{- \ \ t} & non\, attiva \end{cases}, \qquad << \lambda$$

Con il metodo delle fasi è immediato riconoscere che la funzione di distribuzione della Y è una ipoesponenziale a 3 stadi :

exp(3λ+) → exp(3λ) → exp(2λ)

e quindi, ponendo per comodità: $\gamma_1 = 3\lambda +$; $\gamma_2 = 3\lambda$; $\gamma_3 = 2\lambda$

si ottiene:

$$F_Y(t) = \sum_{i=1}^{3} a_i \left(1 - e^{-\gamma_i t}\right) \quad \text{dove } a_i \triangleq \prod_{j \neq i}^{3} \frac{\gamma_j}{\gamma_j - \gamma_i}$$

L' affidabilità del sistema è dunque:

$$R(t) = 1 - F_Y(t) = 1 - \sum_{i=1}^{3} a_i \left(1 - e^{-\gamma_i t}\right)$$

Nella pagina seguente sono raffigurati i grafici con i seguenti valori numerici:

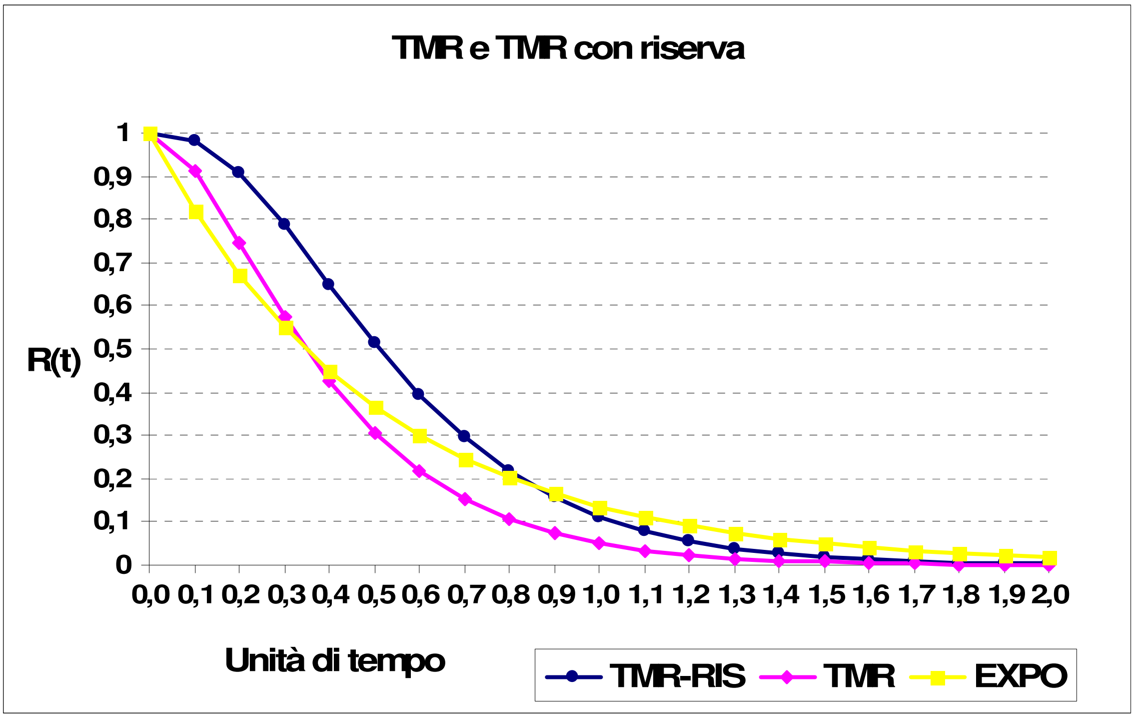
λ = 2	γ1 = 6.2	a1 = 54.5
μ= 0,2	γ2 = 6	a2 = -62
	γ3 = 4	a3 = 8.45

λ = 1	γ1 = 3.2	a1 = 25
μ= 0,2	γ2 = 3	a2 = -32
	γ3 = 2	a3 = 8

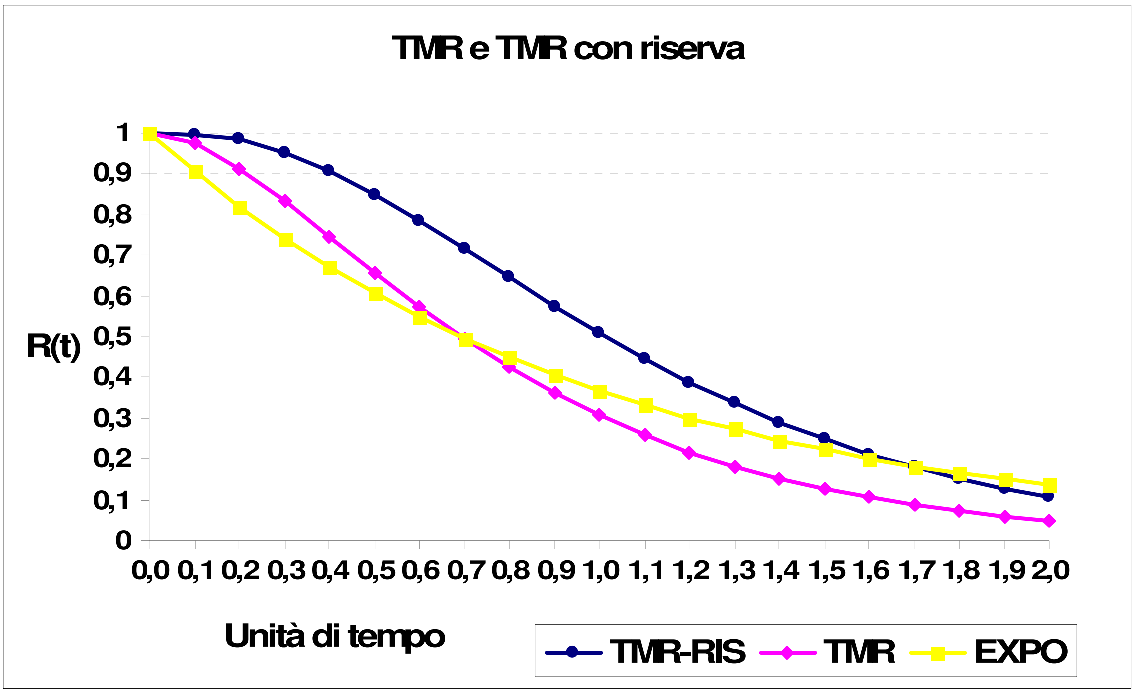
Per esercizio, si provi ad estendere il diagramma delle fasi al caso di commutazione imperfetta.

Confronto di curve di affidabilità

$$\lambda = 2\,, \qquad = 0{,}2 \quad (tempo^{-1})$$



$$\lambda = 1\,, \qquad = 0{,}2 \quad (tempo^{-1})$$



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 $X_i \triangleq$ tempo di arrivo del messaggio i con $i = 1,...,10$ e siano queste variabili aleatorie tutte distribuite secondo la stessa legge esponenziale di parametro $\lambda = \frac{1}{2}\sec^{-1}$ ovvero:

$F_{X_i}(t) = 1 - e^{-\lambda t}$ con $i = 1,...,10$.

Sia $X = X_1 + X_2 + ... + X_{10} \triangleq$ 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: $P(5 \le X \le 10) = F_X(10) - F_X(5)$

$$F_X(10) = 1 - e^{-0,5\cdot 10}\sum_{i=0}^{9}\frac{(0,5\cdot 10)^i}{i!}$$

$$= 1 - 0,007\cdot(1 + 5 + 12,5 + 20,833 + 26,042 + 26,042 + 21,701 + 15,502 + 9,688 + 5,382)$$

$$\cong 0,03183$$

$$F_X(5) = 1 - e^{-0,5\cdot 5} \sum_{i=0}^{9} \frac{(0,5 \cdot 5)^i}{i!} =$$

$$= 1 - 0,82 \cdot (1 + 2,5 + 3,125 + 2,604 + 1,628 + 0,814 + 0,339 + 0,121 + 0,038 + 0,011)$$

$$\cong 0,00028$$

$$F_X(10) - F_X(5) = 0,03183 - 0,00028 = 0,03155$$

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 <u>indipendenti</u>, 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 $X_i \triangleq$ 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 $\lambda_1 = 1/40 m\sec = 25Hz$ e $\lambda_2 = 1/80 m\sec = 12{,}5Hz$, rispettivamente.

Allora $X = X_1 + X_2 \triangleq$ la durata aleatoria del ciclo di attività della CPU è una variabile aleatoria la cui distribuzione si ricava come segue:

$$F_X(t) = P(X_1 + X_2 \le t) = \int_{x_2=0}^{t} F_{X_1}(t - x_2) f_{X_2}(x_2) dx_2 .$$

Dunque

$$F_X(t) = \int_{x_2=0}^{t} \left[1 - e^{-\lambda_1 (t - x_2)}\right] \lambda_2 e^{-\lambda_2 x_2} dx_2 = \lambda_2 \int_{x_2=0}^{t} e^{-\lambda_2 x_2} dx_2 - \lambda_2 e^{-\lambda_1 t} \int_{x_2=0}^{t} e^{-(\lambda_2 - \lambda_1) x_2} dx_2$$

$$= -\frac{\lambda_2}{\lambda_2} e^{-\lambda_2 x_2}\Big|_0^t + \frac{\lambda_2}{\lambda_2 - \lambda_1} e^{-(\lambda_2 - \lambda_1) x_2}\Big|_0^t$$

$$= -\left(e^{-\lambda_2 t} - e^{-\lambda_2 0}\right) + \frac{\lambda_2}{\lambda_2 - \lambda_1} e^{-\lambda_1 t}\left(e^{-(\lambda_2 - \lambda_1)t} - e^{-(\lambda_2 - \lambda_1)0}\right)$$

$$= 1 - e^{-\lambda_2 t} + \frac{\lambda_2}{\lambda_2 - \lambda_1} e^{-\lambda_1 t}\left(e^{-(\lambda_2 - \lambda_1)t} - 1\right) = 1 - e^{-\lambda_2 t} + \frac{\lambda_2}{\lambda_2 - \lambda_1} e^{-\lambda_2 t} - \frac{\lambda_2}{\lambda_2 - \lambda_1} e^{-\lambda_1 t}$$

$$= 1 - e^{-\lambda_2 t}\left(1 - \frac{\lambda_2}{\lambda_2 - \lambda_1}\right) - \frac{\lambda_2}{\lambda_2 - \lambda_1} e^{-\lambda_1 t} = 1 - e^{-\lambda_2 t}\left(\frac{\lambda_2 - \lambda_1 - \lambda_2}{\lambda_2 - \lambda_1}\right) - \frac{\lambda_2}{\lambda_2 - \lambda_1} e^{-\lambda_1 t}$$

$$= 1 + \frac{\lambda_1}{\lambda_2 - \lambda_1} e^{-\lambda_2 t} - \frac{\lambda_2}{\lambda_2 - \lambda_1} e^{-\lambda_1 t} = 1 - \left(\frac{\lambda_2}{\lambda_2 - \lambda_1} e^{-\lambda_1 t} - \frac{\lambda_1}{\lambda_2 - \lambda_1} e^{-\lambda_2 t}\right).$$

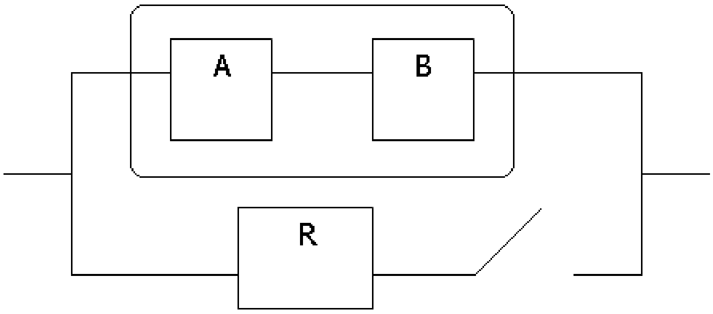
2. la probabilità che la lunghezza di un ciclo non superi i 50msec. è data da:

$$P(X \le 50 m\sec) = F_X(50 m\sec) = 1 - \left(\frac{\lambda_2}{\lambda_2 - \lambda_1} e^{-\lambda_1 t} - \frac{\lambda_1}{\lambda_2 - \lambda_1} e^{-\lambda_2 t}\right)$$

$$= 1 - \left(-1 \cdot e^{-12{,}5 \cdot 0{,}05} + 2 e^{-25 \cdot 0{,}05}\right) = 1 - (-0{,}2865 + 1{,}0705) = 1 - 0{,}784 = 0{,}216$$

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:

   $X_A \triangleq$ tempo al guasto del componente A;

   $X_B \triangleq$ tempo al guasto del componente B;

   $X_R \triangleq$ tempo al guasto del componente di riserva R

   ciascuna distribuita con legge esponenziale di parametro $\lambda_A = \lambda_B = \frac{1}{10000} = 10^{-4}$ e $\lambda_R = \frac{1}{8000} = 1{,}25 * 10^{-4}$, rispettivamente.

   Il tempo medio al guasto del sottosistema seriale è calcolabile a partire dalla formula dell'affidabilità di detto sistema ovvero:

$$R_{AB}(t) = R_A(t) \cdot R_B(t) = \left[1 - F_{X_A}(t)\right] \cdot \left[1 - F_{X_B}(t)\right] = \left[1 - 1 + e^{-\lambda_A t}\right] \cdot \left[1 - 1 + e^{-\lambda_B t}\right] = e^{-(\lambda_A + \lambda_B)t}$$

da cui

$$E[X_{AB}] = \int_{u=0}^{\infty} R_{AB}(u)du = \int_{u=0}^{\infty} e^{-(\lambda_A + \lambda_B)u} du$$

$$= -\frac{1}{\lambda_A + \lambda_B} e^{-(\lambda_A + \lambda_B)u} \Big|_0^{\infty} = -\frac{1}{2 \cdot 10^{-4}} e^{-2 \cdot 10^{-4} \cdot \infty} + \frac{1}{2 \cdot 10^{-4}} e^{-2 \cdot 10^{-4} \cdot 0} = 5000$$

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:

   $Y \triangleq$ 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 assuma che la commutazione riesca sicuramente e si realizzi in tempo trascurabile. Sia

   $Y \triangleq$ tempo al guasto del sistema $= X_{AB} + X_R$.

   Le funzioni di distribuzione e densità sono rispettivamente

   $$\begin{cases} F_{X_{AB}}(t) = 1 - e^{-(\lambda_A + \lambda_B)t} \\ f_{X_{AB}}(t) = (\lambda_A + \lambda_B) e^{-(\lambda_A + \lambda_B)t} \end{cases}$$

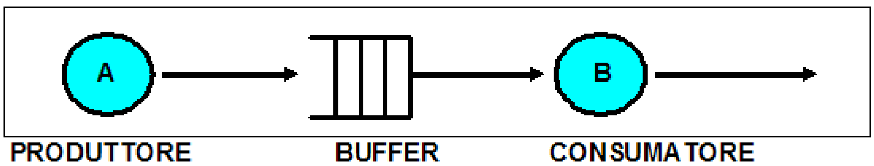
$$\begin{cases} F_{X_R}(t) = 1 - e^{-\lambda_R t} \\ f_{X_R}(t) = \lambda_R e^{-\lambda_R t} \end{cases}$$

Dai risultati ottenuti sulla distribuzione della somma di variabili aleatorie indipendenti, si ha:

$$F_Y(t) = \int_{x_1=0}^{t} F_{X_R}(t - x_1) f_{X_{AB}}(x_1) dx_1 = \int_{x_1=0}^{t} \left(1 - e^{-\lambda_R (t - x_1)}\right)\left[(\lambda_A + \lambda_B) e^{-(\lambda_A + \lambda_B) x_1}\right] dx_1$$

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 casuali, 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 1/ 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.

<u>FASE PRELIMINARE: Il modello di occupazione del buffer</u>

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 ro, 0<ro<1, e un posto libero con probabilità 1-ro. 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:

$$P_n = (1-\rho)\rho^n , \quad n = 0, 1, \ldots, k, k+1, \ldots$$

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* $\triangleq\ n \mid n \le k$

Probabilità evento condizionante: $$P_{n\le k} = \sum_{n=0}^{k} (1-\rho)\rho^n$$

Probabilità evento d'interesse: $$P_{n|n\le k} = \frac{P_n}{P_{n\le k}} = \frac{(1-\rho)\rho^n}{\sum_{n=0}^{k}(1-\rho)\rho^n}, \quad n = 0,1,....,k$$

<u>Derivazione della formula finale per</u> $P_{n|n\le k}$, $n = 0,1,....,k$

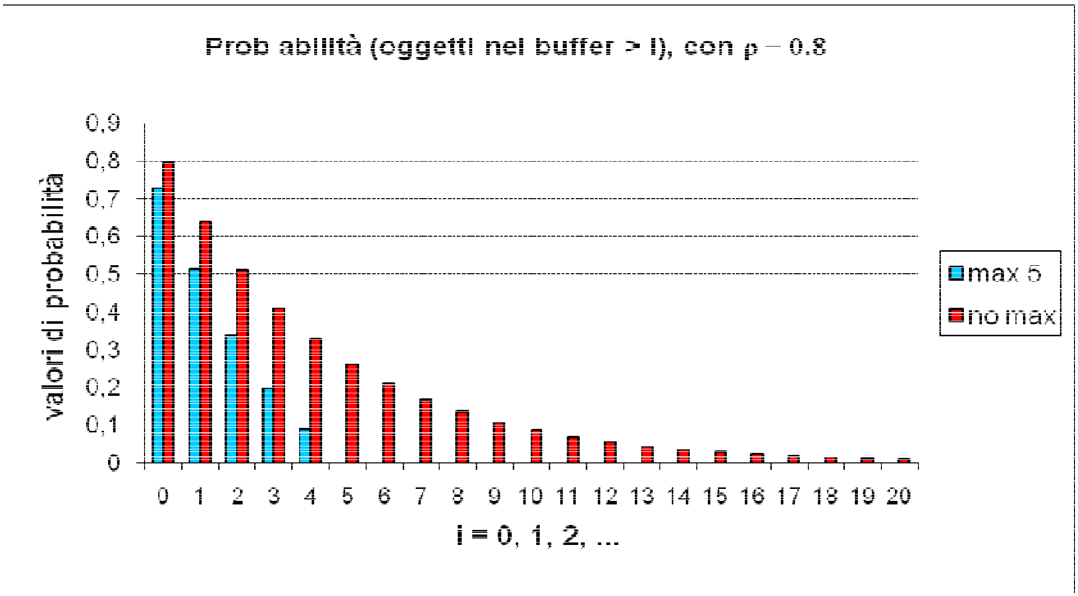
Si sviluppa:

$$\sum_{n=0}^{k} \rho^n = \sum_{i=0}^{\infty} \rho^n - \sum_{n=k+1}^{\infty} \rho^n \qquad con\ \ \rho < 1\ \ e\ \ j \triangleq n-(k+1) \Rightarrow$$

$$= \frac{1}{1-\rho} - \sum_{j=0}^{\infty} \rho^{k+1}\rho^j = \frac{1}{1-\rho} - \frac{\rho^{k+1}}{1-\rho} = \frac{1-\rho^{k+1}}{1-\rho}$$

E si ottiene:

$$P_{n|n\le k} = \frac{\rho^n}{\sum_{n=0}^{k} \rho^n} = \frac{(1-\rho)}{1-\rho^{k+1}}\rho^n, \ n = 0,1,....,k \ \textbf{(Geometrica Troncata)}$$



<u>PRIMA FASE: la distribuzione condizionata del tempo allo svuotamento</u>

Adesso, adottando una legge esponenziale di parametro 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*":

$F_{W|N}(t \mid n) = 1 - e^{-\ t}\sum_{i=0}^{n-1}\frac{(\ t)^i}{i!}$; dove, $F_{W|N=0}(t) = 1$ e $\ \hat{=}\ 1/E[S]$ (ritmo di consumo)

Con $P(N=n) = \dfrac{(1-\rho)}{1-\rho^{k+1}}\rho^n, \quad n = 0,1,\ldots,k \;\Rightarrow\; F_{W|N=0}(t)\cdot P(N=0) = 1\cdot(1-\rho)$

<u>SECONDA FASE: applicazione della formula della distribuzione totale</u>

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 2u.t., assumendo un tempo medio del singolo consumo pari a 1 (E[S]=1).

A titolo di esempio, si adotterà una <u>geometrica troncata</u> 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:

$$F_W^{(m)}(t) \triangleq \Pr[W \le t] = \sum\nolimits_{n=0}^{4} \Pr[W \le t \mid N=n]\cdot \Pr[N=n] = \sum\nolimits_{n=0}^{4} F_{W|N}(t \mid n)\cdot \Pr[N=n]$$

Si ottiene:

$$F_W(t) = \left[1-e^{-\mu t}\right]\cdot P(N=1) + \left[1-e^{-\ t}\cdot\left(1+\frac{(\ t)^1}{1!}\right)\right]\cdot P(N=2) +$$

$$+\left[1-e^{-\ t}\left(1+\frac{(\ t)^1}{1!}+\frac{(\ t)^2}{2!}\right)\right]\cdot P(N=3)\Big]$$

Da cui la probabilità che il tempo di svuotamento superi le 2 u.t.:

$$1-F_X^{(m)}(t=2) \;=\; 1-[\ 0.865\cdot 0.274 + (1-0.135\cdot 3)\cdot 0.206 +$$

$$+\ (1-0.135\cdot 5)\cdot 0.155\ ] = \ 0.59$$

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:

$$E[Y] \triangleq \int_{y=0}^{\infty} y f_Y^{(m)}(y)dy$$

si esprime la densità (marginale) f mediante la formula della densità totale:

$$= \int_{y=0}^{\infty} y \int_{x=0}^{\infty} f_{Y|X}(y/x) f_X^{(m)}(x)dx \; dy \quad = \int_{x=0}^{\infty} \int_{y=0}^{\infty} y f_{Y|X}(y/x)dy \;\; f_X^{(m)}(x)dx$$

e riconoscendo che è naturale definire

come <u>*valore atteso condizionato*</u> il seguente: $E[Y \mid X] \triangleq \int_{y=0}^{\infty} y f_{Y|X}(y/x)dy$

si ottiene la <u>*formula del valore atteso totale*</u>:

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

Con una variabile condizionante, X, discreta:

$$E[Y] = \sum_{x=0}^{\infty} E[Y \mid X] P_X(x)$$

<u>OSSERVAZIONE</u>:

La formula del valore atteso condizionato non definisce un numero, ma una funzione sullo spazio delle realizzazioni (continue e non negative) della X:

$$E[Y \mid X = x], \qquad 0 \le x < \infty .$$

Tale funzione, che descrive l'andamento del valore atteso della Y al variare della x, è detta <u>*curva di regressione*</u> e si può ricavare, in linea di principio, a partire dalla conoscenza della densità condizionata, $f_{Y|X}(y \mid x)$ (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:

$$E[(X - E[X])(Y - E[Y])] \triangleq COV(X,Y)$$

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:

$$\begin{aligned} COV(X,Y) &\triangleq E[(X - E[X])(Y - E[Y])] \\ &= E[XY - YE[X] - XE[Y] + E[X]E[Y]] \\ &= E[XY] - E[X]E[Y] \end{aligned}$$

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 <u>non correlate</u>.

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 <u>indipendenza</u> o solo di <u>correlazione nulla</u>, si noti che risulta valida la seguente implicazione sulla varianza della somma di X e Y:

$$\begin{aligned} VAR[X + Y] &= VAR[X] + VAR[Y] + 2\,COV(X,Y) && (*) \\ &= VAR[X] + VAR[Y] && \text{(additività della varianza)} \end{aligned}$$

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: $-1 \le \rho(X,Y) \le 1$ 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: $-1 \le \rho(X,Y) \le 1$

Prova:

Siano W una varia aleatoria $E[W^2] \neq 0$ e Z una variabile aleatoria $E[Z^2] \neq 0$

allora

$$(aW - Z)^2 \ge 0 \quad (a \in \Re) \Rightarrow E[(aW - Z)] \ge 0.$$

Ma

$$E\left[(aW - Z)^2\right] = a^2 E\left[W^2\right] - 2a \cdot E[W \cdot Z] + E\left[Z^2\right] \ge 0.$$

Definendo

$$a \triangleq E[W \cdot Z] / E\left[W^2\right]$$

per sostituzione si ottiene

$$\frac{(E[W \cdot Z])^2}{\left(E\left[W^2\right]\right)^2} E\left[W^2\right] - 2a \cdot \frac{(E[W \cdot Z])^2}{E\left[W^2\right]} + E\left[Z^2\right] \geq 0$$

$$\Rightarrow -\frac{(E[W \cdot Z])^2}{E\left[W^2\right]} + E\left[Z^2\right] \geq 0 \quad \Rightarrow \quad \frac{(E[W \cdot Z])^2}{E\left[W^2\right] \cdot E\left[Z^2\right]} \leq 1 .$$

Con

$$W \hat{=} X - {}_{X} \text{ e } Z \hat{=} Y - {}_{Y} \Rightarrow -1 \leq \rho(X, Y) \leq +1 .$$

<u>Proprietà 2</u>

Si dimostra che: $|\rho(X, Y)| = 1 \quad \Leftrightarrow \quad Y = \alpha + \beta X$,con α e β reali e β ≠ 0

Prova: (si dimostra solo l'implicazione ⇐)

Dalla relazione: $Y = \alpha + \beta X$

si ha:

$$\begin{aligned} COV[X, Y] &= E[XY] - E[X]E[Y] \\ &= \alpha E[X] + \beta E\left[X^2\right] - E[X](\alpha + \beta E[X]) \\ &= \beta VAR[X] \end{aligned}$$

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:

$$E[Y \mid X = x] \triangleq a \cdot x + b, \qquad 0 \le x < \infty$$

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.

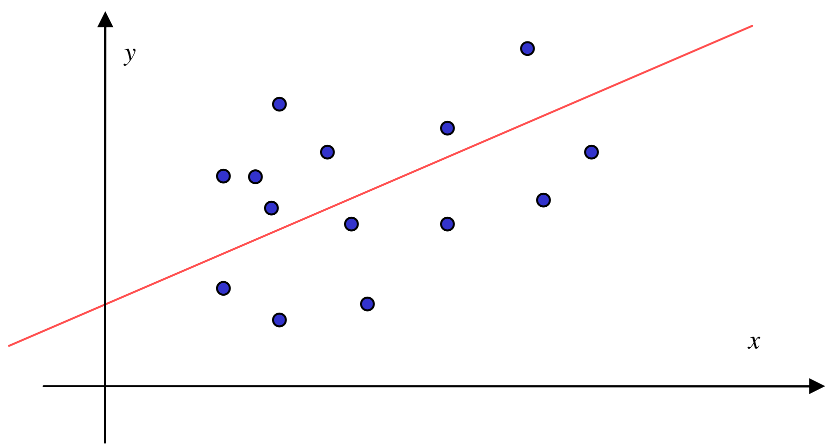
<u>TEOREMA</u> (sui coefficienti della retta di regressione)
Quale che sia la forma delle funzioni di distribuzione della X e della Y, i coefficienti a e b possono essere ricavati dalle seguenti formule:

$$a = \rho \frac{\sqrt{VAR[Y]}}{\sqrt{VAR[X]}} = \frac{COV(X,Y)}{VAR[X]}; \qquad b = E[Y] - \frac{COV(X,Y)}{VAR[X]} E[X];$$

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 $y_1, y_2, \ldots, y_n$ osservati rispetto ai valori sulla retta stessa: S = $\sum_{i=1}^{n} \left[y_i - (ax_i + b) \right]^2$.

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:

$$\begin{cases} \dfrac{\partial S(a,b)}{\partial b} = -2 \cdot \sum_{i=1}^{n} (y_i - b - ax_i) = 0 \\ \dfrac{\partial S(a,b)}{\partial a} = -2 \cdot \sum_{i=1}^{n} (y_i - b - ax_i) \cdot x_i = 0 \end{cases} \quad ; \quad \text{da cui:}$$

$$\begin{cases} n \cdot b + a \cdot \sum_{i=1}^{n} x_i - \sum_{i=1}^{n} y_i = 0 \\ b \cdot \sum_{i=1}^{n} x_i + a \cdot \sum_{i=1}^{n} x_i^2 - \sum_{i=1}^{n} x_i y_i = 0 \end{cases}$$

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 <u>risultano</u> <u>in</u> <u>accordo</u> con le formule già ricavate: $a = \dfrac{COV(X,Y)}{VAR[X]}$ e $b = E[Y] - a \cdot E[X]$.

<u>Esempi di rette di regressione</u>

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.

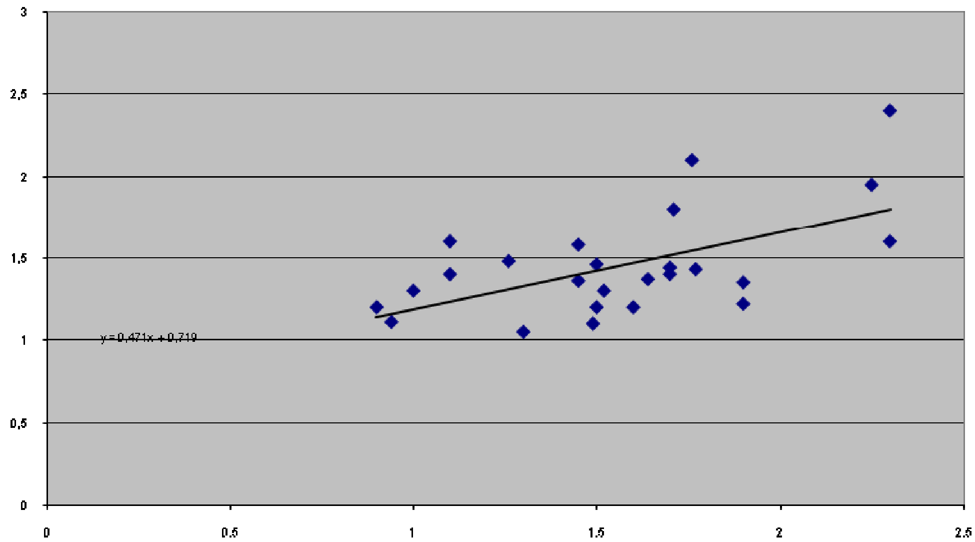


Fig. 1. Retta di regressione per i tempi di giacenza (settimane)

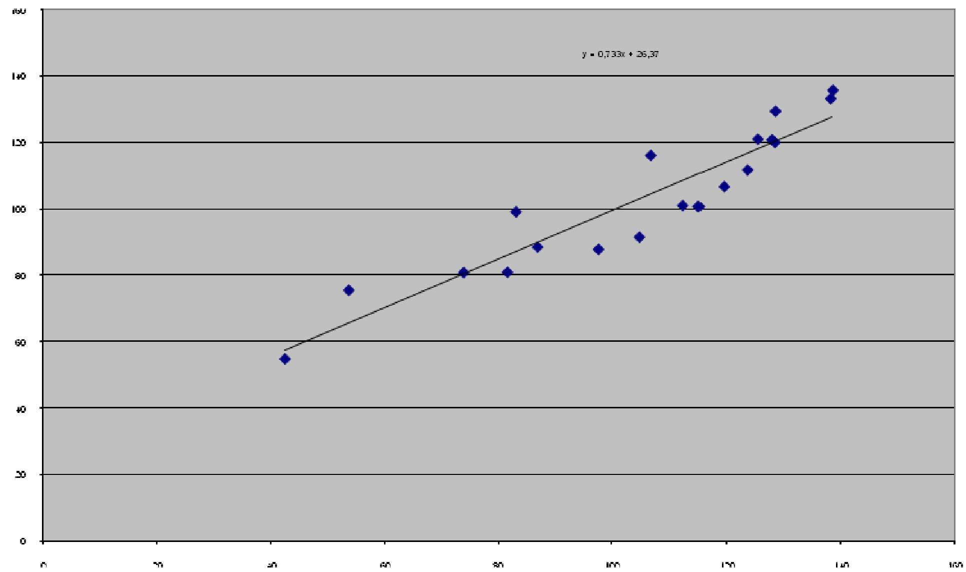


Fig. 2. Retta di regressione per i livelli di occupazione (unità)

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)

$$y = 0{,}4717x + 0{,}7194$$

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

$$y = 0{,}733x + 26{,}37$$

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 $\quad = 1$, 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:

$$\rho = a \cdot \frac{\sqrt{VAR[X]}}{\sqrt{VAR[Y]}}$$,   con X e Y variabili aleatorie generiche.

Allora, riprendendo la formula di stima di "a", ricavata in precedenza:

$$\hat{a} = \frac{n \cdot \sum_{i=1}^{n} x_i y_i - \sum_{i=1}^{n} x_i \sum_{i=1}^{n} y_i}{n \cdot \sum_{i=1}^{n} x_i^2 - \left(\sum_{i=1}^{n} x_i \right)^2} \quad \text{e riscrivendola così:} \quad \hat{a} = \frac{\left(n \cdot \sum_{i=1}^{n} x_i y_i - \sum_{i=1}^{n} x_i \sum_{i=1}^{n} y_i \right) \Big/ n^2}{\left[n \cdot \sum_{i=1}^{n} x_i^2 - \left(\sum_{i=1}^{n} x_i \right)^2 \right] \Big/ n^2}$$

si perviene alle seguenti ($\bar{x}$ e $\bar{y}$ indicano le medie aritmetiche degli *n* rispettivi valori)

$$\hat{a} = \frac{\frac{1}{n}\sum_{i=1}^{n} x_i y_i - \bar{x} \cdot \bar{y}}{\frac{1}{n}\sum_{i=1}^{n} x_i^2 - \bar{x}^2} = \frac{\frac{1}{n}\sum_{i=1}^{n} \left[(x_i - \bar{x}) \cdot (y_i - \bar{y}) \right]}{\frac{1}{n}\sum_{i=1}^{n} (x_i - \bar{x})^2} \qquad (+)$$

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:

$\hat{\rho}(P,C) = 7.37 \quad 10^{-5}$, in risposta al punto 1)

$\hat{\rho}(V,G) = 0.993$, in risposta al punto 2)

$\hat{\rho}(P_{[1971-2000]}, G) = -0.433$ e $\hat{\rho}(P_{[2001-2030]}, G) = -0.095$, 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. Vice versa, 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

$$\rho \triangleq \frac{COV(X,Y)}{\sqrt{VAR[X] \cdot VAR[Y]}}$$

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?