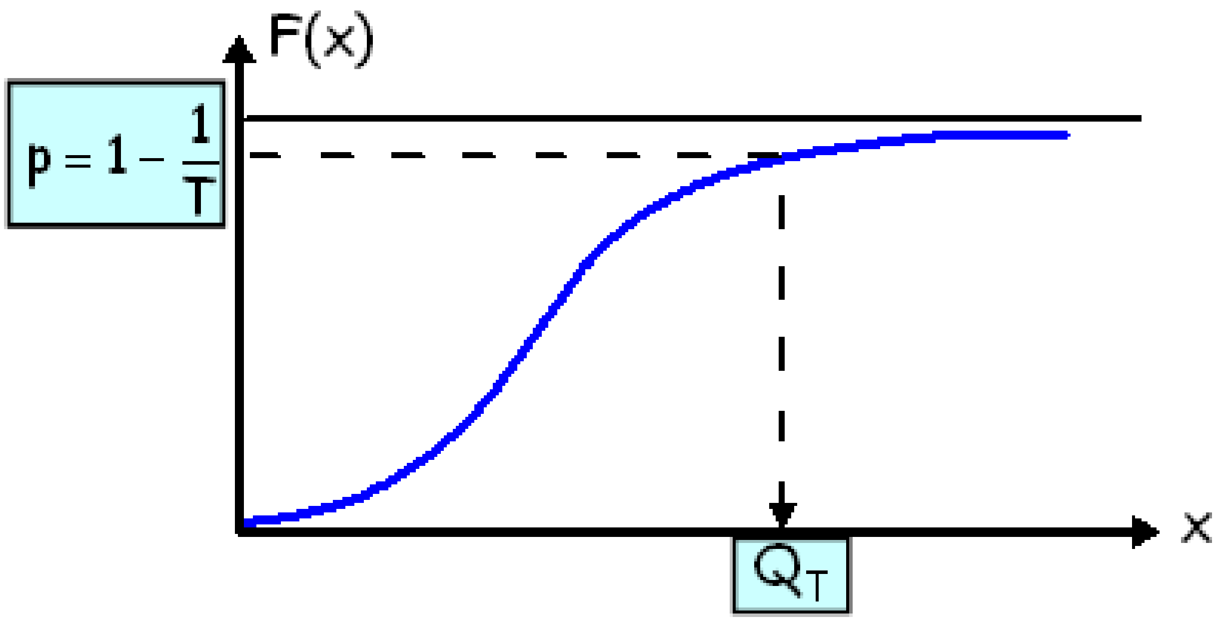
Idraulica e Idrologia: Lezione 7

Agenda del giorno

- **Uso delle distribuzioni statistiche;**
- **Distribuzione di Gumbel;**
- **Carte probabilistiche;**
- **Plotting position;**
- **Elaborazioni delle precipitazioni estreme;**
- **Curve di probabilità pluviometrica;**
- **Calcolo della Linea Segnalatrice di Probabilità Pluviometrica tramite Excel.**

Quantile

Definizione:
Il quantile relativo alla probabilità p è il valore caratterizzato da una probabilità di non superamento pari a p. Nota che p può essere collegata al tempo di ritorno T da p=1-(1/T)



Indicazioni sul tempo di ritorno

Volendo dare qualche indicazione intorno al tempo di ritorno da adottare in problemi applicativi, si possono assumere gli orientamenti riportati in tabella.

Tipo di opera	tempo di ritorno
Ponti e difese spondali	100-150
Difese dei torrenti	30-50
Dighe	500-1000
Bonifiche	15-25
Fognature urbane	5-10
Tombini e ponticelli di corsi d'acqua	30-50
Sottopassi stradali	50-100
Cunette e fossi di guardia	10-20

Parametri statistici

L'obiettivo della statistica è quello di estrarre l'informazione essenziale da un insieme di dati sintetizzandolo in un certo numero di parametri.

popolazione | **campione**

$$\mu = \int_{-\infty}^{+\infty} x f(x)dx \qquad tendenza\ centrale\ (media) \qquad \bar{x} = \frac{1}{n}\sum_{i=1}^{n} x_i$$

$$\sigma^2 = \int_{-\infty}^{+\infty} (x-\mu)^2 f(x)dx \qquad variabilita'\ \ (varianza) \qquad s^2 = \frac{1}{n-1}\sum_{i=1}^{n}(x_i-\bar{x})^2$$

$$\gamma = \frac{\int_{-\infty}^{+\infty}(x-\mu)^3 f(x)dx}{\sigma^3} \qquad simmetria\ (coeff.di\ asimmetria) \qquad C_s = \frac{n\sum_{i=1}^{n}(x_i-\bar{x})^3}{(n-1)(n-2)s^3}$$

… …

Misure di dispersione: varianza e deviazione standard

La varianza, come pure la deviazione standard (la radice quadrata della varianza) costituisce una misura classica di dispersione. Come la media, risulta sensibile alla presenza di outliers.

$$\sigma^2 = \int_{-\infty}^{+\infty} (x-\mu)^2 f(x)dx \quad dispersione\,(\text{varianza})\; s^2 = \frac{1}{n-1}\sum_{i=1}^{n}(x_i - \bar{x})^2$$

Deviazione standard = s

$$s = \sqrt[2]{s^2}$$

Misure di dispersione: deviazione standard e coefficiente di variazione

Una misura di dispersione frequentemente utilizzata è costituita dal coefficiente di variazione (**CV**), definito come rapporto fra la deviazione standard campionaria **s** e la media campionaria **m**.

$$CV = \frac{s}{m}$$

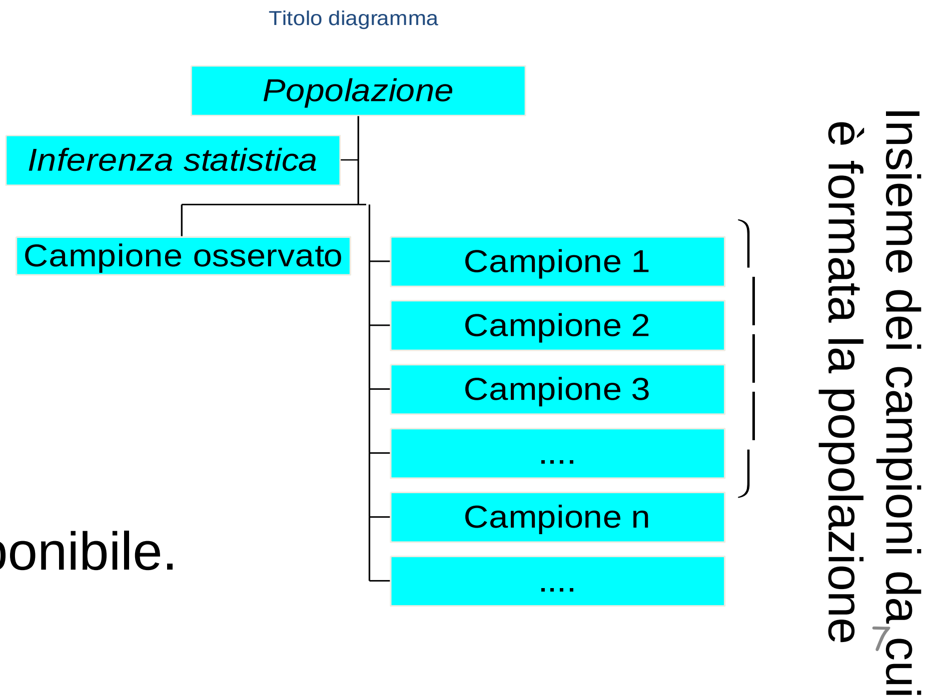
Il coefficiente di variazione risulta definito solo se la media è diversa da zero. Viene utilizzato soprattutto per confrontare la dispersione di campioni diversi.

Trattamento probabilistico dei dati idrologici

Come collegare la magnitudo degli eventi estremi alla loro frequenza di accadimento tramite l'impiego di distribuzioni di probabilità?

L'insieme dei dati disponibili (p. es. 30 valori di portate massime annuali osservate presso una certa sezione di chiusura di un bacino) viene considerato come un campione estratto da una ipotetica popolazione di dimensione infinita.

Le proprietà statistiche dei campioni (media, varianza,..) variano da campione a campione, mentre quelle della popolazione sono uniche.
Tramite tecniche di ***inferenza statistica***, le proprietà della popolazione vengono stimate a partire da quelle del campione disponibile.



Uso di una funzione di distribuzione di probabilità per descrivere un campione di dati

Processo di inferenza statistica

Dato un campione di dati $x_1, x_2, x_3,.., x_n$, della variabile casuale X, ci si pone il problema di determinare la forma di una funzione F(x) atta a rappresentare, con ragionevole approssimazione, la distribuzione vera, ma incognita, della X.

Processo di inferenza statistica

- scelta della distribuzione
  In base a ragionamenti sulla naturale variabilità di X si presuppongono una o più forme analitiche per la distribuzione incognita
- stima dei parametri della distribuzione
  metodo dei momenti
  metodo della massima verosimiglianza
  metodi grafici
- test di controllo
  viene valuata l'affidabilità delle distribuzioni ipotizzate
  test di Kolmogorov-Smirnov
  test del Chi-quadro

Stima dei parametri della distribuzione

Una funzione di distribuzione di probabilità rimane un'astrazione fino a quando non viene collegata alle osservazioni del fenomeno fisico. Questo si ottiene stimando i parametri della distribuzione sulla base dei dati del campione disponibile.

Tecniche disponibili:

- **metodo dei momenti**
- **metodo della massima verosimiglianza**
- **metodi grafici**

Metodo dei momenti

- I parametri delle distribuzioni sono esprimibili in funzione dei parametri statistici della popolazione (media, varianza,..)

- Si assume che i parametri statistici della popolazione coincidano con quelli del campione

- I parametri della distribuzione prescelta vengono espressi in funzione dei parametri statistici del campione

Distribuzione di probabilità utilizzata nel corso

Distribuzioni:

Gumbel

Distribuzione di Gumbel - 1

Funzione di densità di probabilità

$$f(x) = \frac{1}{\alpha}\exp\left[-\frac{x-u}{\alpha} - \exp\left(-\frac{x-u}{\alpha}\right)\right]$$

$$F(x) = \exp\left[-\exp\left(-\frac{x-u}{\alpha}\right)\right]$$

Le funzioni possono essere espresse utilizzando la variabile ridotta w=(x-u)/α

I due parametri sono: α e u

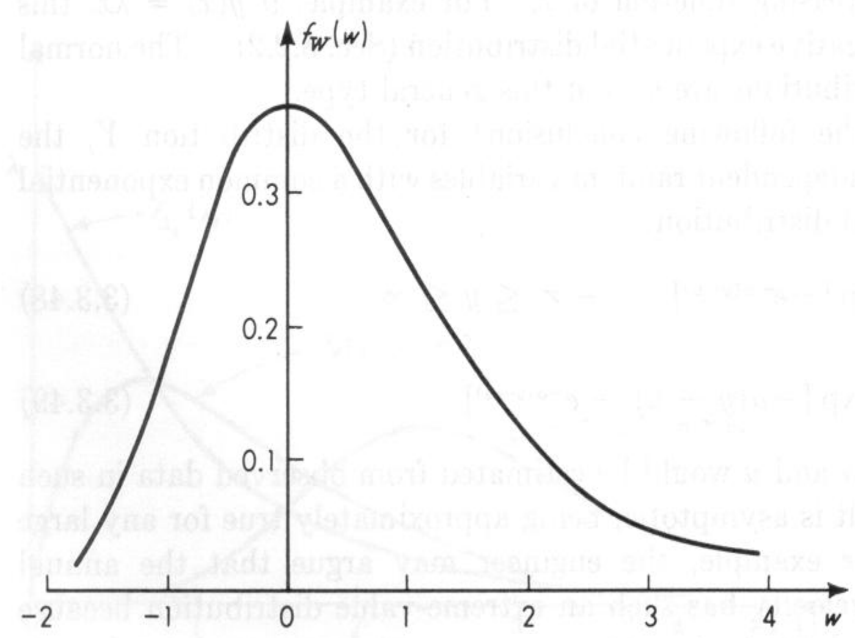
Metodo dei momenti per stima di α e u ⇓

$$\alpha = \frac{\sqrt{6}s}{\pi}$$

$$u = \bar{x} - 0.5772\alpha$$

Distrib. di densità ⇒

di probabilità di Gumbe



w=(x-u)/α

Distribuzione di Gumbel - 2

Determinazione del valore di x caratterizzato da un tempo di ritorno T

$$F(x) = \exp[-\exp(-w)]$$

$$\Rightarrow w = -\ln\left[\ln\left(\frac{1}{F(x)}\right)\right]$$

poiché è :

$$\frac{1}{T} = P(X \geq x_T) = 1 - P(X \leq x_T) = 1 - F(x_T)$$

$$\Rightarrow F(x_T) = \frac{T-1}{T}$$

$$w_T = -\ln\left[\ln\left(\frac{T}{T-1}\right)\right]$$

$$x_T = u + \alpha w_T$$

Distribuzione di Gumbel - 3

Determinazione del valore di T corrispondente ad un'osservazione x

Noto il valore di x ed dei parametri u e α , si calcola il valore di w, (variabile ridotta), da cui si calcola F(X≤x) e quindi il valore del tempo di ritorno T

$$
\begin{aligned}
& w = \frac{x-u}{\alpha} \\
& \Rightarrow F(x) = \exp[-\exp(-w)] \\
& \Rightarrow T = \frac{1}{1-F(x)}
\end{aligned}
$$

Esempio di applicazione: distribuzione di Gumbel

Stazione pluviografica di Trento
principali statistici del campione e valori dei parametri della distribuzione di probabilità di GUMBEL
Stima dei parametri u e α con il metodo dei momenti.

Durata (ore)	N casi	minimo	massimo	media	Deviazione standard	u	α
1	52	11.0	45.0	21.1	7.7	17.69288	6.00279
3	52	16.0	62.0	31.3	11.0	26.35318	8.59703
6	52	25.0	82.4	42.5	14.4	35.98954	11.23940
12	52	30.0	123.0	58.6	21.8	48.77891	17.02172
24	52	40.2	147.6	76.8	25.5	65.31335	19.89397

Valori dei quantili regolarizzati di precipitazione (mm) con la distribuzione di probabilità di GUMBEL

T	w
2	0.36651
5	1.49994
20	2.97020
50	3.90194
100	4.60015

	Tempo di ritorno (T)				
Durata	**2**	**5**	**20**	**50**	**100**
1	19.9	26.7	35.5	41.1	45.3
3	29.5	39.2	51.9	59.9	65.9
6	40.1	52.8	69.4	79.8	87.7
12	55.0	74.3	99.3	115.2	127.1
24	72.6	95.2	124.4	142.9	156.8

Esercizio: calcolare con il metodo di Gumbel la pioggia oraria di tempo di ritorno ventennale per la stazione di Lago Verde

Stazione Lago Verde (PAB)
Valori pioggia massima annuale
Durata: 1 ora

18.4
8.6
9.0
11.2
13.4
10.4
10.4
13.0
16.2
10.4
15.4
8.4
9.8
23.6
18.0

Media=13.08 mm
Dev St=4.41 mm

$$\alpha = \frac{\sqrt{6}s}{\pi} = 3.44$$

$$u = \bar{x} - 0.5772\alpha = 11.09$$

$$F(x) = \exp[-\exp(-w)]$$

$$\Rightarrow w = -\ln\left[\ln\left(\frac{1}{F(x)}\right)\right]$$

poiché è:

$$\frac{1}{T} = P(X \geq x_T) = 1 - P(X \leq x_T) = 1 - F(x_T)$$

$$\Rightarrow F(x_T) = \frac{T-1}{T}$$

$$w_T = -\ln\left[\ln\left(\frac{T}{T-1}\right)\right]$$

$$x_T = u + \alpha w_T$$

$$w_{20} = 2.97$$

$$\Rightarrow x_{20} = 11.09 + 3.44 \cdot 2.97 = 21.3\text{mm}$$

EJ Gumbel: matematico ed oppositore al regime nazista

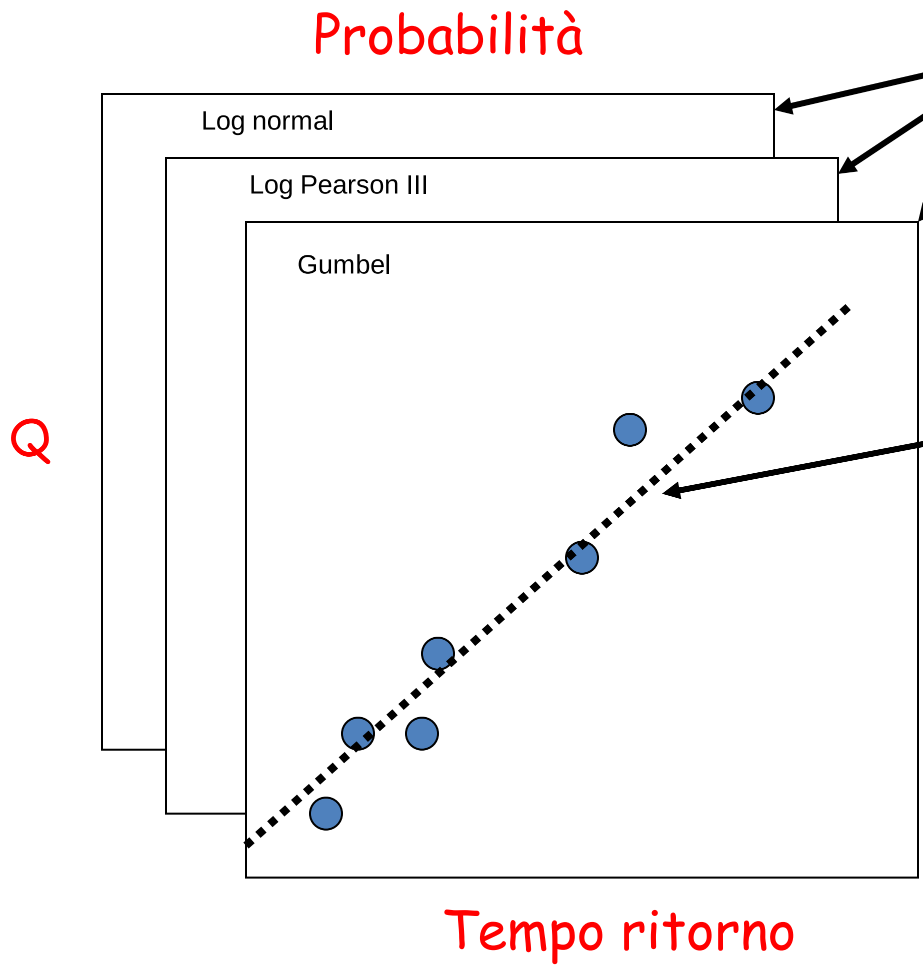
EJ Gumbel nasce a Monaco il 18 luglio 1891 e muore a New York il 10 settembre 1966. EJ Gumbel non è noto solo per la sua attività di ricerca nel campo della matematica, ma anche per la sua attività come pacifista ed oppositore al regime nazionalsocialista. Gumbel inizia la sua attività di ricerca ad Heidelberg nel 1923. Tuttavia, le sue critiche alle attività del partito nazista lo costringono a riparare in Francia nel 1932, ed infine negli USA dal 1940. Viene accolto all'università di Standford, e nel 1958 pubblica la sua opera più famosa: 'Statistics of extremes', dove appare anche la distribuzione che da lui prende nome. Nell'introduzione a questo libro, EJ Gumbel scrive:
"This book is written in the hope, contrary to expectation, that humanity may profit by even a small contribution to the progress of science".

Le carte probabilistiche

Le **carte probabilistiche** sono specifiche per ogni tipo di funzione di probabilità (log-normale, Gumbel, ..) e vengono costruite in modo tale che le curve di probabilità della
funzione corrispondente vi vengono rappresentate da rette.
⇒ le carte probabilistiche possono essere utilizzate per verificare l'ammissibilità della funzione
di probabilità prescelta per descrivere il campione, ancor prima di stimare i parametri:
se il tipo di funzione di distribuzione prescelto è adatto ad interpretare le osservazioni, i punti devono addensarsi intorno ad una retta.

Su queste carte, normalmente in ascissa viene riportata la variabile x, mentre in ordinata viene
riportata la probabilità $P(X \geq x)$ o $P(X \leq x)$, il tempo di ritorno, o la variabile ridotta della distribuzione.

Metodo Grafico



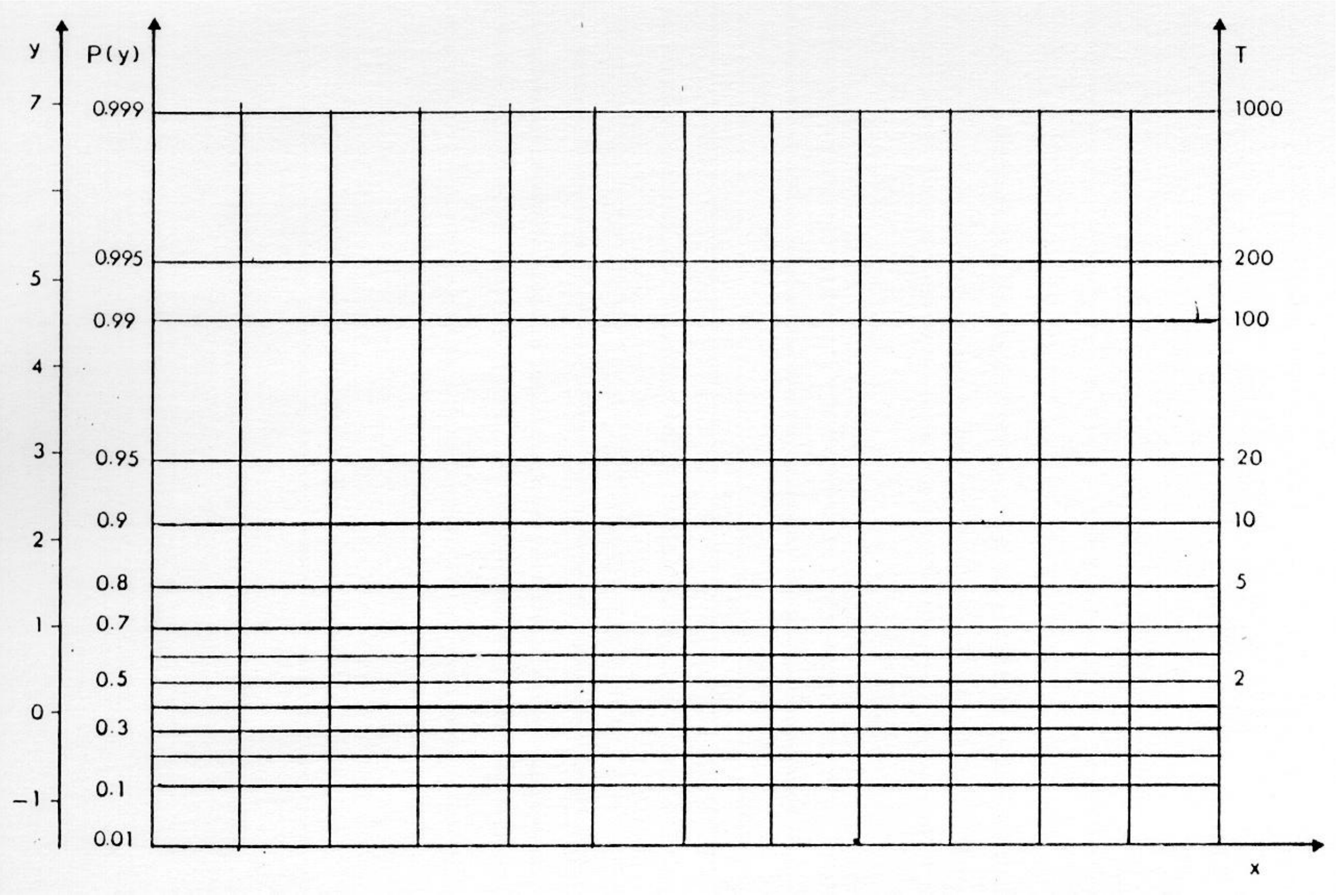
Ciascuna distribuzione di probabilità ha la propria carta probabilistica

I dati vengono riportati in ciascuna carta, dove vengono anche interpolati linearmente

Viene selezionata la distribuzione cui corrisponde la carta probabilistica dove i punti si dispongono con minimo scarto intorno alla retta regolarizzatrice.

La carta probabilistica di Gumbel

Carta probabilistica di Gumbel



Carte probabilistiche e plotting position

PLOTTING POSITION

Per riportare un punto sulla carta probabilistica, è necessario conoscere di esso il valore x e la probabilità $P(X \geq x)$ o $P(X \leq x)$. Se i parametri della distribuzione non sono stati ancora determinati, non è possibile calcolare la probabilità attraverso le formule consuete. Nelle carte probabilistiche viene quindi utilizzata un'approssimazione della probabilità di superamento, detta ***plotting position***.

Approssimazione normalmente utilizzata:

$$\boxed{P(X \geq x) = \frac{m}{N+1}}$$

dove

m : posizione del dato nella serie ordinata in senso decrescente,

N : numerosità del campione

Dopo aver riportato i valori sulla carta probabilistica, ed avere accertato che si addensano intorno ad una retta, è possibile:

- **utilizzare direttamente il diagramma per identificare la retta che meglio regolarizza i valori (p.es. tramite il metodo dei minimi quadrati)**

oppure

- **procedere in modo analitico alla determinazione dei parametri (p. es. tramite il metodo dei momenti) e quindi riportare la retta risultante sul grafico al fine di valutarne la capacità descrittiva (*<u>questo metodo è preferibile</u>*).**

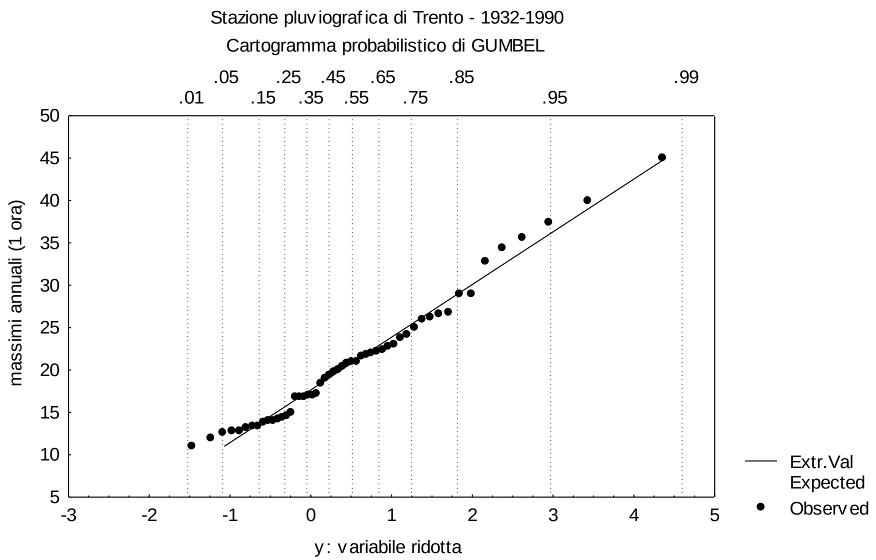
Applicazione: stazione pluviografica di Trento

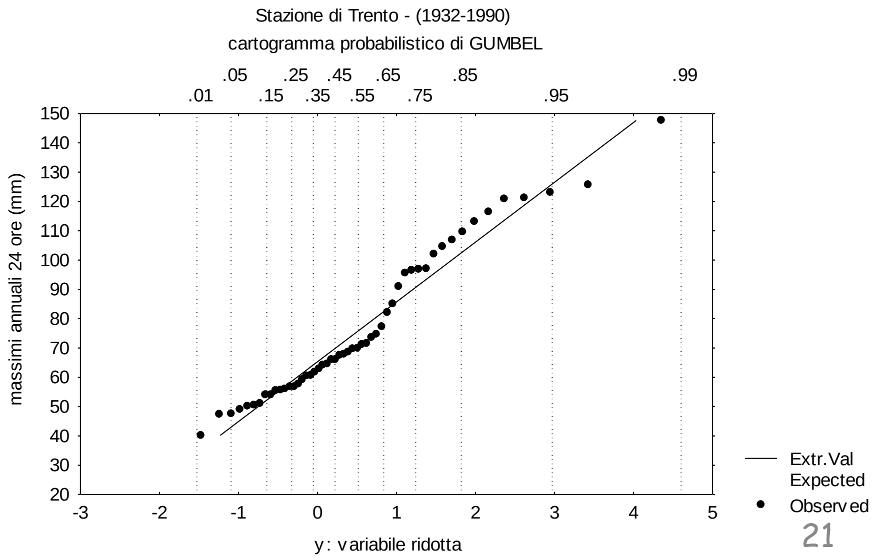
Le figure riportano le carte probabilistiche di Gumbel relative ai valori di precipitazione massima annuale di durata pari ad 1 e 24 ore.

Le rette introdotte nei grafici sono quelle
stimate tramite il metodo dei momenti. Infatti nelle carte viene riportata in ascissa la variabile ridotta ed in ordinata il dato.

$$\boxed{y = \frac{x-u}{\alpha} \Rightarrow x = y\alpha + u}$$

P. es.: u rappresenta il valore dell'intercetta
dall'applicazione del metodo dei momenti si ricava
u (1h)   = 17.7
u (24h) = 65.3





Elaborazione delle precipitazioni - 1

Premessa - 1

Le tecniche di elaborazione statistica-probabilistica descritte alle sezioni precedenti possono essere applicate sia per l'analisi delle serie di valori massimi annuali di portata che di precipitazione. Generalmente, l'obiettivo dell'analisi è costituito dall'individuazione del valore di portata al picco caratterizzato da un assegnato tempo di ritorno (quantile di portata).

Il modo più diretto per conseguire questo risultato è rappresentato dall'elaborazione delle serie di dati di portata massima annuale (per una certa sezione idrometrica). Tuttavia, la distribuzione delle stazioni di misura idrometrica sul territorio, per quanto estesa possa essere la rete di monitoraggio, non è tale da coprire il fabbisogno di conoscenza che si ha intorno a queste grandezze; ovvero, non sempre (anzi, quasi mai) la sezione fluviale o torrentizia presso la quale si attende il valore del quantile, è corredata da una stazione idrometrica.

Elaborazione delle precipitazioni - 2

Premessa - 2

L'esigenza di poter stimare i quantili di interesse in modo generale, ha stimolato quindi l'affermarsi di una serie di metodologie e di modelli atti a consentire la stima della grandezza di interesse (il quantile di portata) a partire dalle piogge. Il processo di modellazione che si segue in tal caso, è quello relativo alla trasformazione afflussi-deflussi.

Lo studio delle precipitazioni estreme (massime annuali) è quindi il primo passo per poter utilizzare efficacemente tali modelli.

La prima osservazione da farsi a tal punto è la seguente:
A parità di rarità dell'evento di pioggia, l'altezza totale di precipitazione non cresce proporzionalmente al crescere della durata.
Questo significa che l'intensità dell'evento decresce al crescere della durata.
Infatti, a parità di rarità, eventi brevi sono generati da fenomeni convettivi (intensi), mentre gli eventi più prolungati sono generati da fenomeni stratiformi (meno intensi).

Elaborazione delle precipitazioni - 3

Perché ci interessa studiare la relazione delle precipitazioni estreme con la durata?

Perché bacini idrografici di estensione diversa entrano in risonanza (ovvero: producono un idrogramma di picco massimo – detto critico) in corrispondenza di durate di precipitazioni diverse: bacini più estesi sono sollecitati criticamente da precipitazioni prolungate nel tempo, mentre bacini più piccoli hanno una durata critica minore.

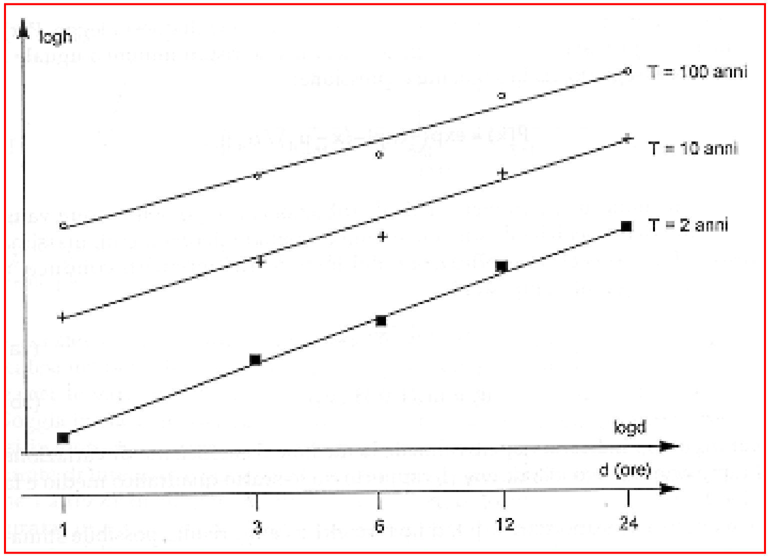
Per disporre di uno strumento *generale* è quindi necessario disporre di una relazione localmente *generale* fra piogge estreme e loro durata.

Curve di probabilità pluviometrica

La curva che fornisce la relazione tra durata *t* e altezza di precipitazione *h* con tempo di ritorno assegnato T prende il nome di *curva di probabilità pluviometrica*

(a volte si usa anche il termine di linea segnalatrice di probabilità pluviometrica, oppure:
curva segnalatrice di possibilità climatica
curva di possibilità climatica
curva segnalatrice di possibilità pluviometrica)

Costruzione della curva segnalatrice di possibilità climatica (è rappresentata nella figura in un piano logaritmico).



Metodo dell'analisi dei massimi annuali. Legge probabilistica adoperata: Gumbel o lognormale.

$$h = at^n$$

h = altezza di pioggia tempo ritorno T

t = durata di pioggia

a, n = parametri

Il tempo di ritorno convenzionale della curva è quello comune ai singoli punti adoperati per costruirla.

DETERMINAZIONE DELLA CURVA DI PROBABILITA' PLUVIOMETRICA

Stima dei parametri *a* ed *n* della curva di probabilità pluviometrica

- Si determinano i quantili di precipitazione per le diverse durate (1, 3, 6, 12 e 24 ore, se l'interesse è per le durate orarie) per i diversi tempi di ritorno di interesse

- Si considerano i quantili relativi ad un assegnato tempo di ritorno e si regolarizzano mediante la relazione

$$\boxed{h = at^{n}}$$

dove a ed n sono due parametri che dipendono dal tempo di ritorno (il valore di n è compreso fra 0 ed 1).

DETERMINAZIONE DELLA LINEA SEGNALATRICE DI PROBABILITA' PLUVIOMETRICA

Stima dei parametri *a* ed *n* della linea segnalatrice di probabilità pluviometrica

Il metodo più rapido per determinare i parametri *a* ed *n* di ciascuna linea segnalatrice consiste nell'interpolarli linearmente in un piano in scala logaritmica (ovvero: in un piano in cui in ascisse vi sia *log t* ed in ordinata *log h*). Infatti:

$$\boxed{h = at^{n} \rightarrow \log h = \log a + n \log t}$$

A tal fine, sia:
x=log t,
y=log h,
b=log a,
per cui $\rightarrow$

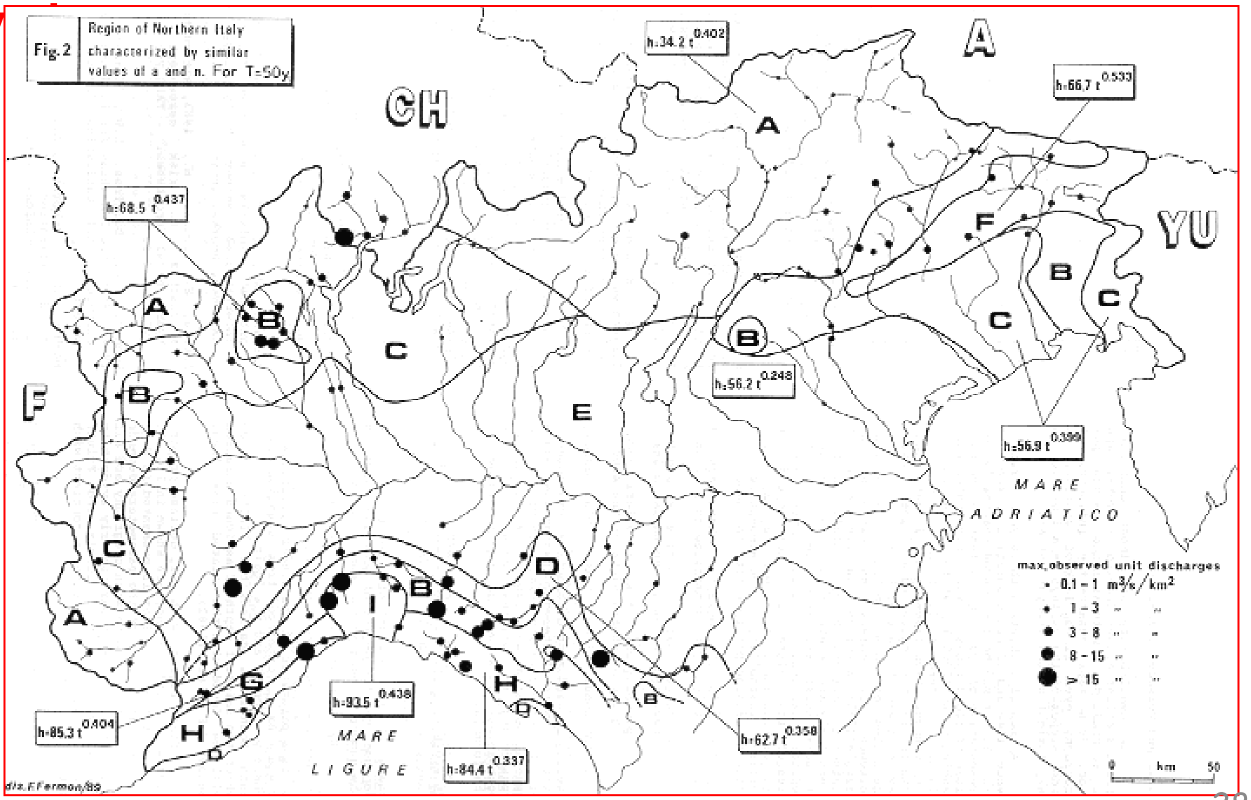
$$n = \frac{m\sum_{i=1}^{m} x_i y_i - \sum_{i=1}^{m} x_i \sum_{i=1}^{m} y_i}{m\sum_{i=1}^{m} x_i^{\,2} - \left(\sum_{i=1}^{m} x_i\right)^2}$$

$$b = \frac{\sum_{i=1}^{m} x_i^{\,2} \sum_{i=1}^{m} y_i - \sum_{i=1}^{m} x_i \sum_{i=1}^{m} x_i y_i}{m\sum_{i=1}^{m} x_i^{\,2} - \left(\sum_{i=1}^{m} x_i\right)^2}$$

$$dove\ m = numero\ di\ \ punti\ utilizzati\ \ (5,\ normalmente)$$

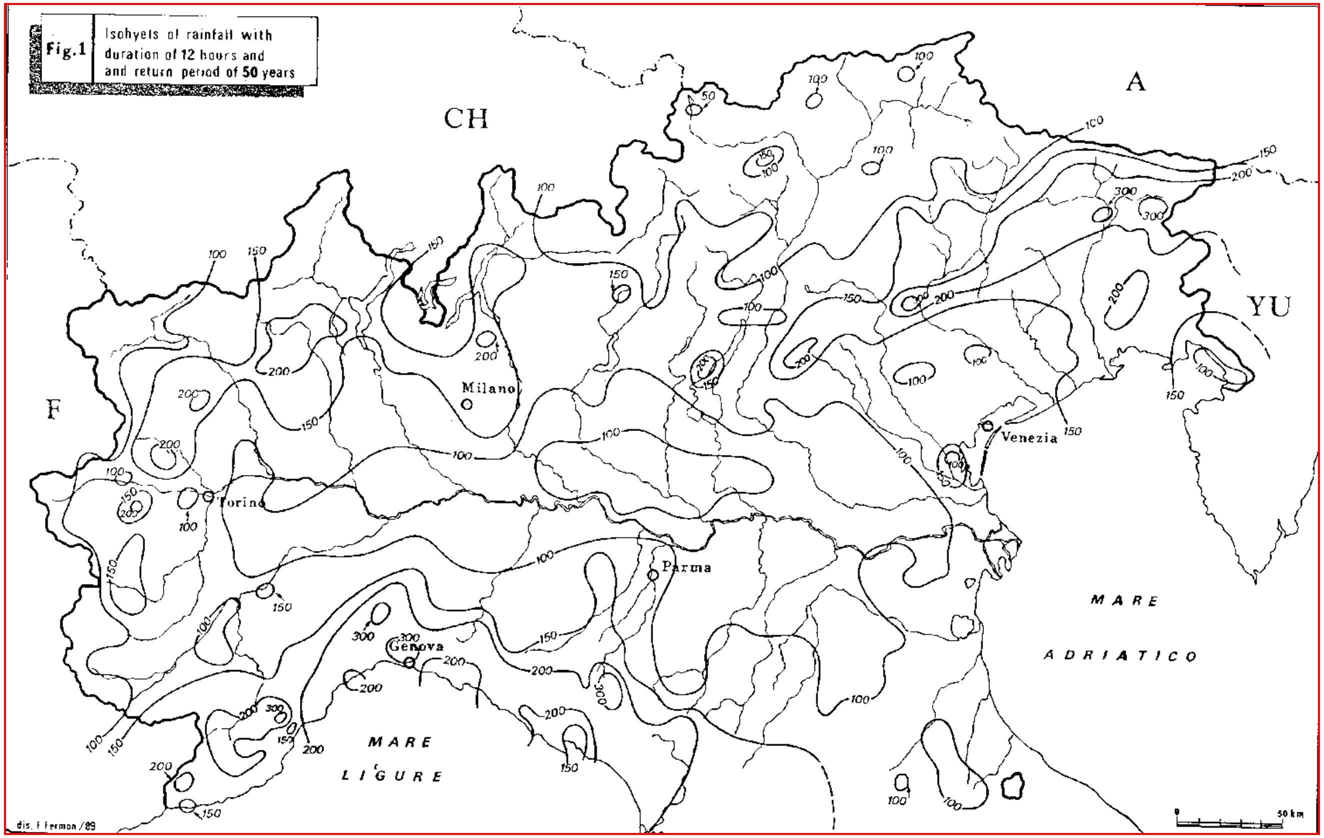
DISTRIBUZIONE SPAZIALE DELLE CARATTETISTICHE DELLE PIOGGE IN ITALIA SETTENTRIONALE

Linee segnalatrici per alcune regioni dell'Italia Settentrionale per tempo di ritorno pari a 50 anni e valori massimi osservati di portata specifica



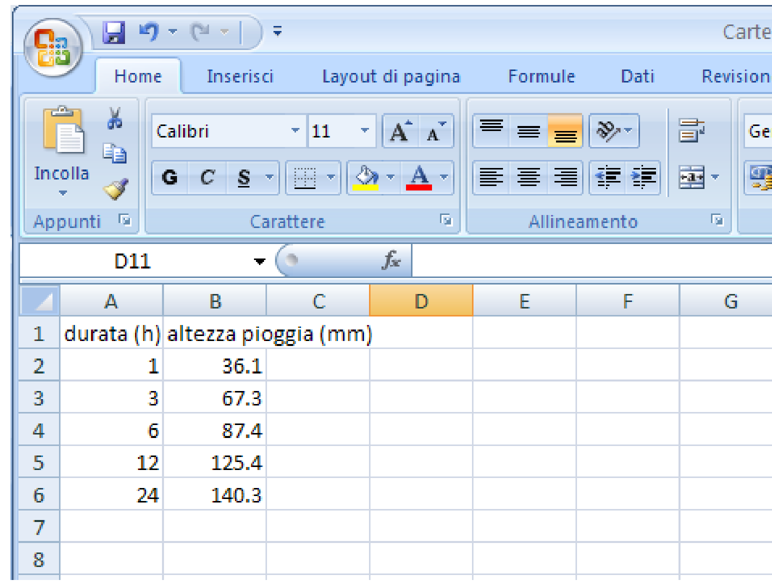
DISTRIBUZIONE SPAZIALE DELLE CARATTERISTICHE DELLE PIOGGE IN ITALIA SETTENTRIONALE

Isoiete delle precipitazioni di durata pari a 12 ore e tempo di ritorno pari a 50 anni

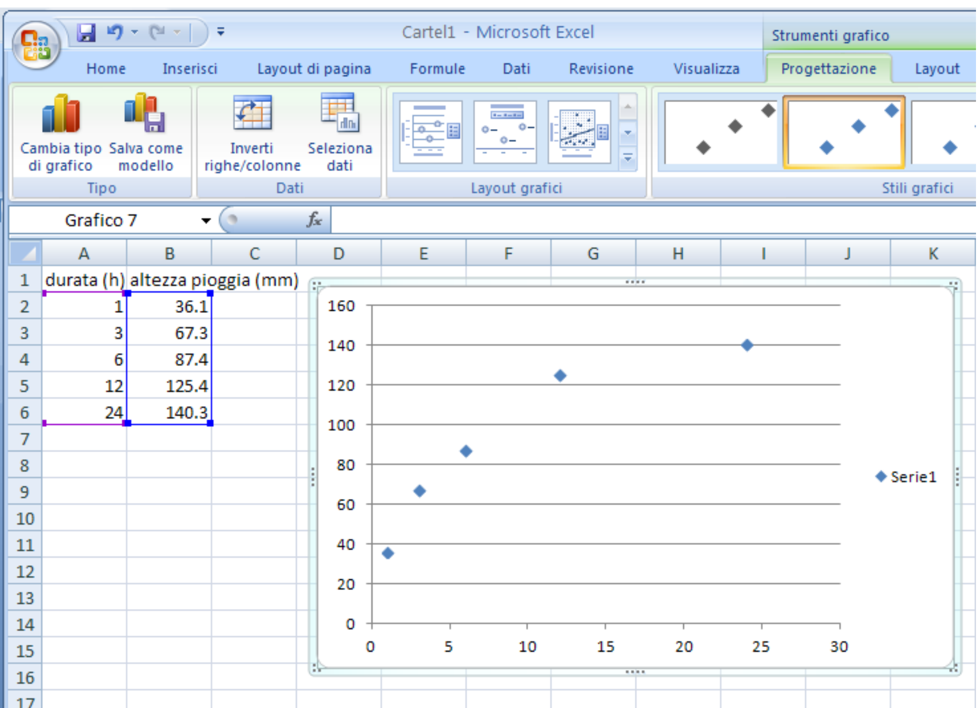


CALCOLO PARAMETRI LINEA SEGNALATRICE CON EXCEL - 1

T	$X_{T\text{-}1h}$	$X_{T\text{-}3h}$	$X_{T\text{-}6h}$	$X_{T\text{-}12h}$	$X_{T\text{-}24h}$
2	18	28	38	54	66
5	23	39	51	73	86
10	26	45	60	85	99
50	33	60	79	113	128
100	36	67	87	125	140
200	39	73	95	136	152

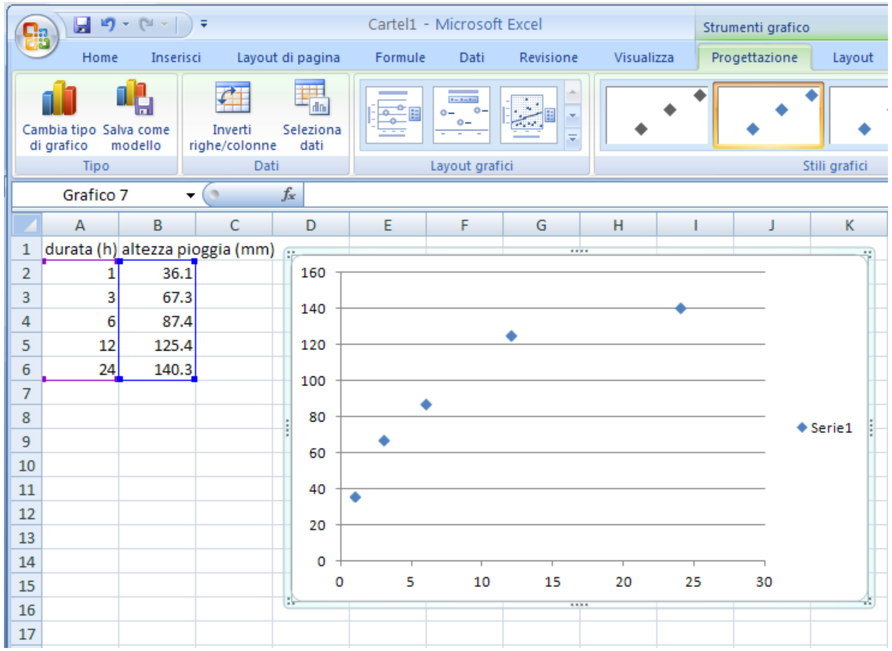


Inserisco i dati in excel

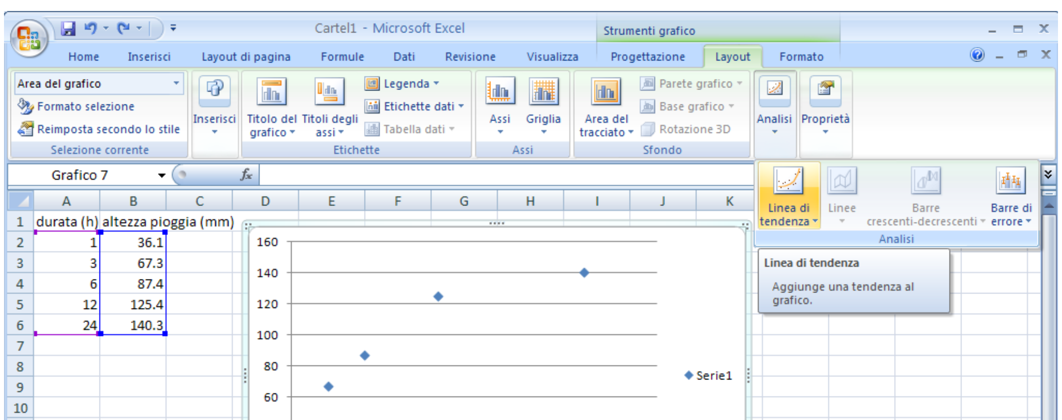


Visualizzo i dati con 'Inserisci' e quindi 'Grafico a dispersione'

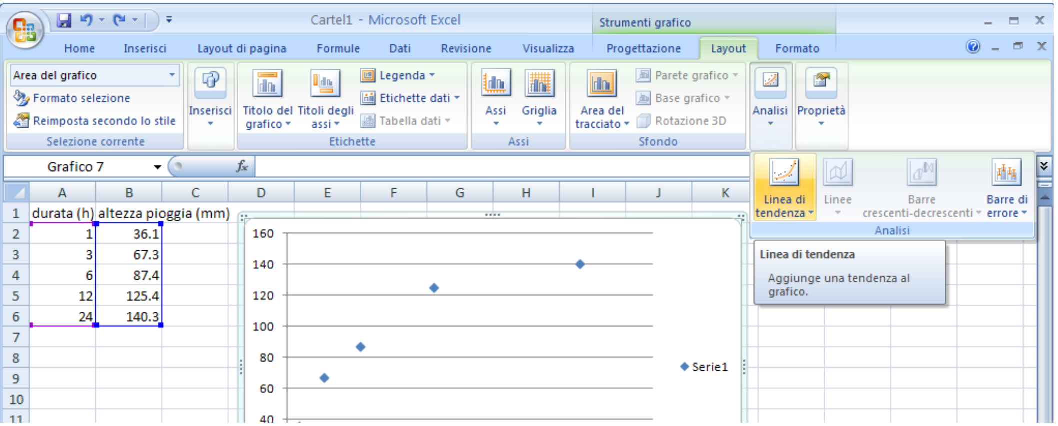
CALCOLO PARAMETRI LINEA SEGNALATRICE CON EXCEL - 2

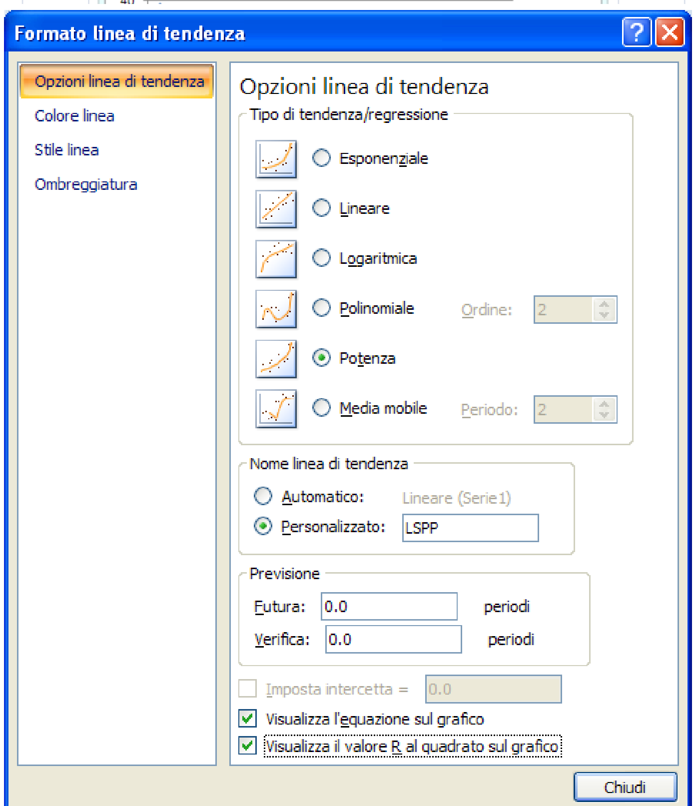


Dopo aver marcato l'area del grafico, vado in 'Layout' e seleziono 'Analisi' dove, successivamente, seleziono 'Linea di tendenza'.



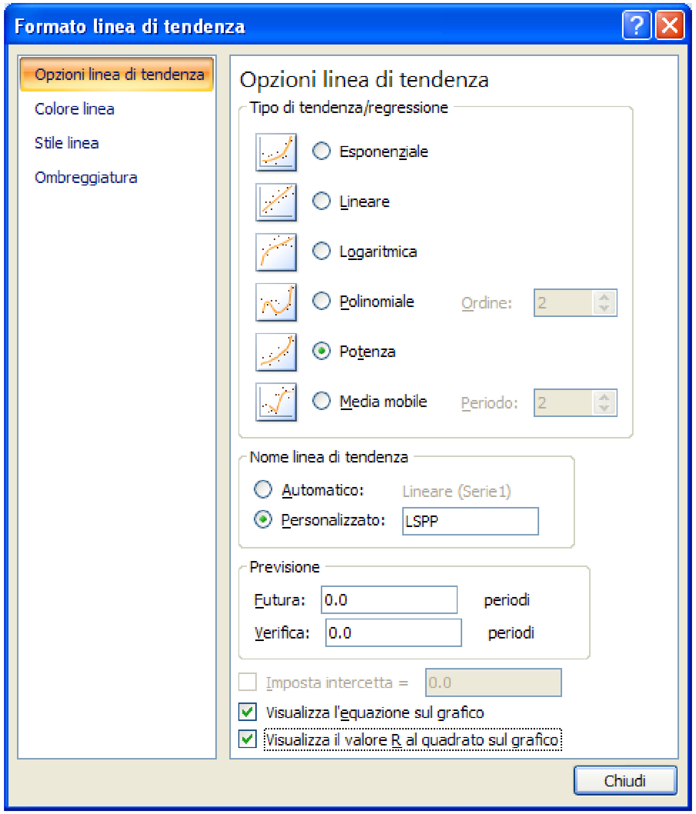
CALCOLO PARAMETRI LINEA SEGNALATRICE CON EXCEL - 3



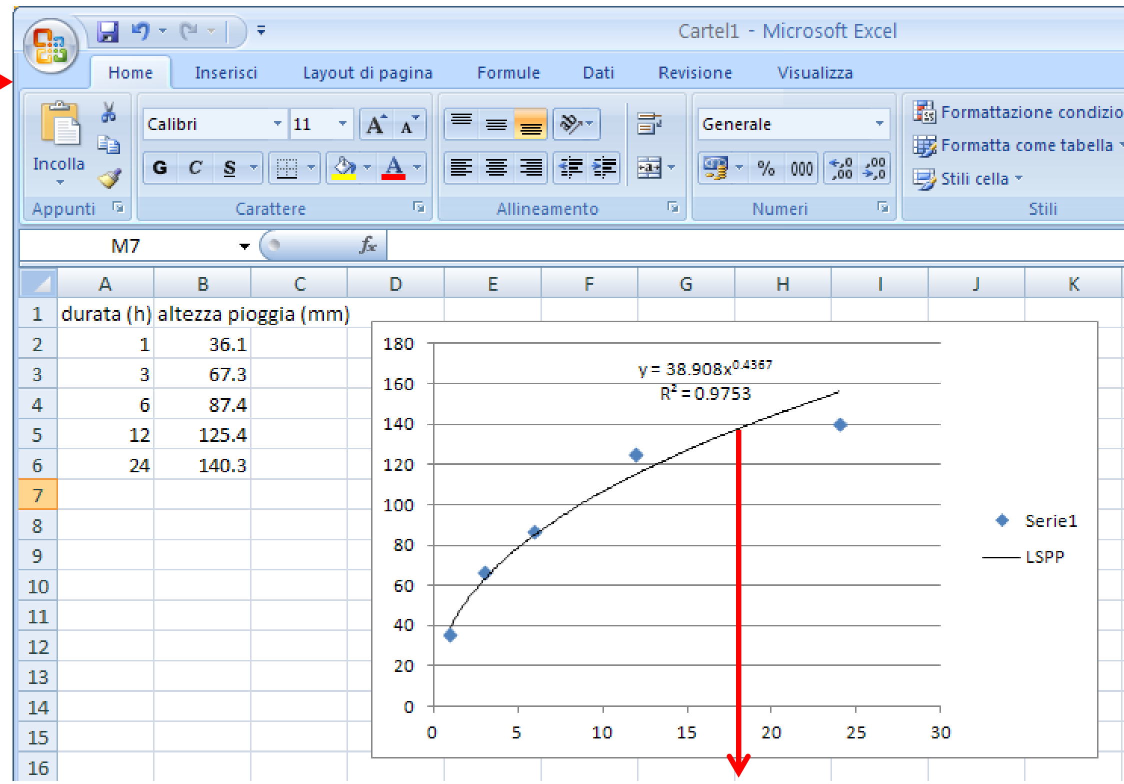


Seleziono il formato di funzione interpolatrice desiderata ('Potenza') e richiedo di visualizzare sia l'equazione sul grafico che il valore di R2.

CALCOLO PARAMETRI LINEA SEGNALATRICE CON EXCEL - 4



L'equazione così calcolata rappresenta la LSPP



Il grafico può essere importato nel documento word della relazione, ma è necessario qualificare le variabili rappresentate in asse x ed y ed aggiungere anche le unità di misura.

$$h = at^n$$

$$h = 38.908t^{0.4367}$$