NUMERO: 811 DATA: 04/02/2014

A P P U N T I

STUDENTE: Ottina

MATERIA: Idrologia 2013/14

Prof. Laio

IDROLOGIA

ACQUA COME RISORSA

L'uomo necessita dell'acqua e la usa per vari scopi: domestico, agricolo, industriale, produzione energia.
Come quantificare la risorsa idrica e come dimensionare le opere?
L'acqua può essere piovana o prelevata dai canali, fiumi e serbatoi (quando è in abbondanza).

ACQUA COME ELEMENTO DI RISCHIO

Eventi alluvionali di piena producono gli straripamenti, in seguito a forti precipitazioni: bisogna progettare le opere per la sicurezza idraulica del territorio.
Bisogna quantificare l'elemento che produce rischio, come il livello del corso d'acqua (per evitare esondazioni) e costruire gli argini.

BISOGNA QUANTIFICARE L'ACQUA.

PORTATA DI UN CORSO D'ACQUA Q $[m^3/s]$

Volume di acqua che passa nella sezione di un corso d'acqua nell'unità di tempo.

Il nostro problema è DIMENSIONARE le opere, come le condotte di una fognatura per la raccolta del drenaggio urbano (uniamo acque bianche e nere).

Le opere possono essere piccole (grondaie) o grosse (fiumi) dove le dimensioni dipendono dall'intensità delle piogge.

© Proprietà riservata dell'autore - Digitalizzazione e distribuzione a cura del CENTRO APPUNTI - Corso Luigi Einaudi, 55 - Torino / Pagina 3 di 202

La portata Q è differente da precipitazione o temperatura perché è una VARIABILE AGGREGATA: l'acqua si accumula nella sezione di chiusura. Infatti la portata non è funzione del punto MA devo sapere quant'acqua appartiene al bacino idrografico.

- PLUVIOMETRO: misura la pioggia nel punto;
- PORTATA, non è sufficiente l'informazione puntuale MA bisogna considerare tutta l'area che appartiene allo spartiacque superficiale.

Se pensiamo ad una rete, bisogna capire qual è l'area complessiva che raccoglie l'acqua che interessa la rete.

In città gli spartiacque sono i tetti delle case

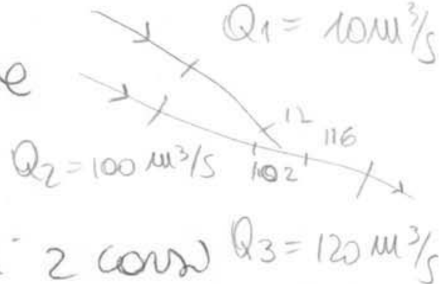
FOGNATURE → ETTARI　　　ARGINI, PONTI → km^2

Perché variabile aggregata?

- PIOGGIA: dalle tre misure avrò che $x \simeq 40 \div 60$ mm, tramite una INTERPOLAZIONE SPAZIALE.

Questo si può fare perché la pioggia è funzione del PUNTO.

- PORTATA: non posso interpolare perché Q, oltre ad essere funzione del punto, è anche funzione del bacino idrografico.
Fino al punto di giunzione di 2 corsi d'acqua posso avere un lieve aumento di portata, dovuto ad altri ingressi.

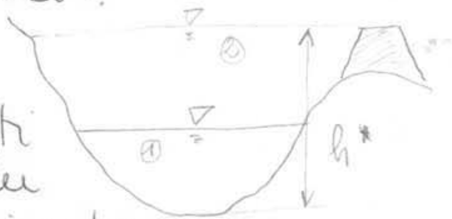


Le portate non possono essere interpolate perché punti vicini possono avere valori significativamente diversi.

SEZIONE CON DISPONIBILITA' DI DATI IDROMETRICI

le sezioni idrometriche vengono messe nei grandi corsi d'acqua e, per utilizzare il dato, non è sufficiente cadere perfettamente nella sezione MA essere nelle vicinanze (± 5 km).

cosa fare con i dati disponibili?



considero un corso d'acqua normale che ha subito straripamenti NON molto gravi: bisogna mettere in sicurezza il corso d'acqua mediante GABBIONATA (nei corsi piccoli / medi) che fa:

- rende la riva meno erodibile;
- evita l'esondazione;
- permette di avere livelli idrici maggiori.

L'opera consente di far passare una maggiore portata senza avere straripamenti.

h^* = <u>ALTEZZA DEL RILEVATO ARGINALE</u>, misurata rispetto al fondo alveo.

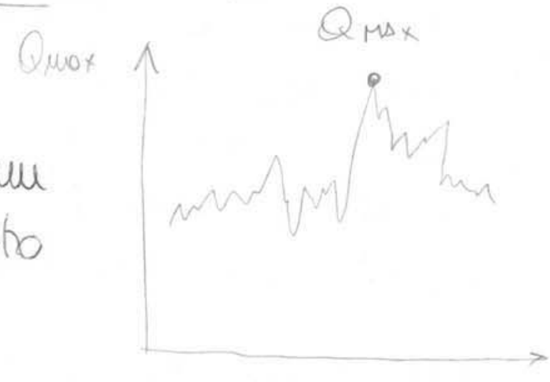
Prima devo determinare h^*, poi la lunghezza e larghezza del rilevato.

Ipotizzo di trovare i dati di portata in una sezione vicina.
Per proteggere il territorio faccio riferimento alle <u>PORTATE MASSIME AL COLMO DI PIENA</u> (Q_{max}).

Ragioniamo su 1 anno.

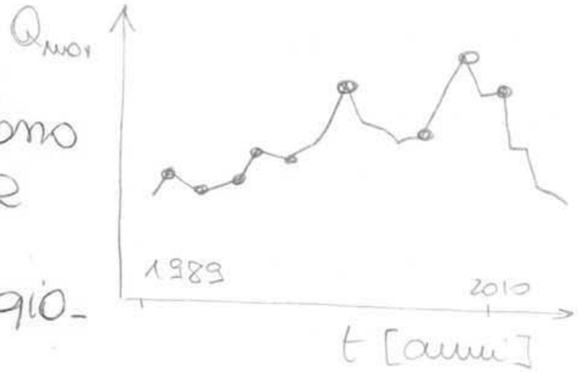
La portata, per sua natura, è un dato continuo nel tempo: riporto l'andamento nel tempo di Q.

Posso avere il dato per molti anni.



IDROGRAMMA 1 ANNO

La portata è una VARIABILE STATISTICA CASUALE e non posso garantire una protezione totale perché potrà sempre succedere qualcosa di peggio.



Q = VARIABILE CASUALE AMBIENTALE ILLIMITATA SUPERIORMENTE

Come dimensionare?
Il dimensionamento non è deterministico MA bisogna accettare un LIVELLO DI RISCHIO: dimensionare l'opera per poter funzionare 95÷99 anni su 100.

- LIVELLO DI RISCHIO

Bisogna accettare che una volta ogni 100 anni può capitare l'alluvionamento anche in presenza dell'opera: esiste un livello di rischio RESIDUO.

Mettiamo l'opera per ridurre il livello di rischio residuo.

ARGINE SI / NO — IL FIUME ESONDA 1 VOLTA ∀ 100 / ∀ 10 ANNI

Qual è il livello di rischio accettabile?

ESEMPIO

Accettiamo che ci sia mediamente (mediamente perché la portata è una variabile casuale e non ho la certezza assoluta) ci sia una sormonta arginale ogni 100 anni.

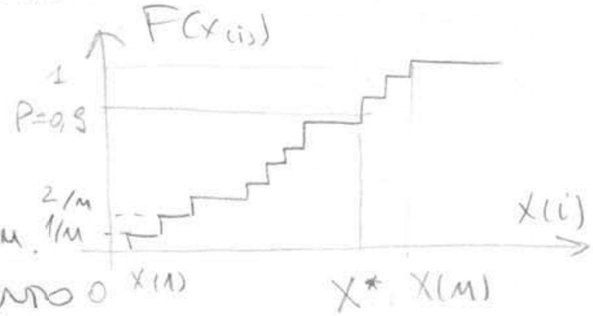
Il valore 1 ∀ 10/100 anni non compete al progettista, perché l'ingegnere conosce Q_{max} ed i livelli di rischio (dalle norme).

- FUNZIONE DI FREQUENZA EMPIRICA

$$\boxed{F(x_{(i)}) = \frac{i}{n}}$$

X = generica variabile casuale
X(i) = i-esimo valore nel campione ordinato
campione: insieme dei punti disponibili per l'inferenza statistica; ordinato in senso crescente
n = dimensione campionaria
$F(x_{(i)})$ rappresenta una frequenza empirica di non superamento, frequenza con la quale un certo valore non viene superato.

Se $x_{(i)} < x_{(1)}$, $F(x) = 0$
Ogni gradino vale $1/n$.
$X_{(n)}$ = massimo assoluto del campione ordinato per $i = n$.
$F(x_n) = 1$ valore di non superamento



Nell'esempio di $R = 1/10$, con $n = 22$ ho che $F(x) = 0,9$ e noto $F(x)$ posso ricavare la $x(i)$.

- PROBABILITA' DI NON SUPERAMENTO $\boxed{P = 1 - R}$

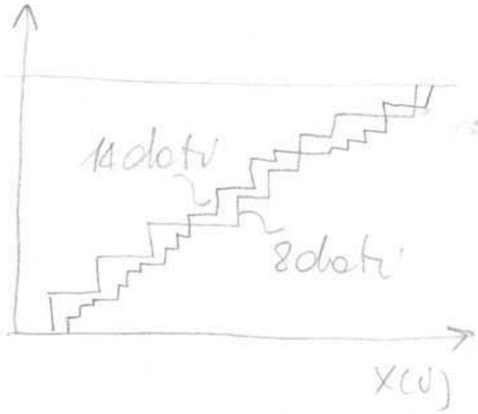
Se il livello di rischio (R) aumenta, la probabilità di non superamento (P, probabilità cumulata), CALA.

Se confondo $P \simeq F(x)$, entro con $P = 0,9$, esco con x^*; MA confondendo P con F(x) faccio l'errore di confondere il campione con la popolazione da cui è stato estratto.

CAMPIONE: insieme finito dei dati realmente disponibili (dimensione n-FINITA)
POPOLAZIONE: insieme infinito dei dati da cui estraggo il valore/campione (la variabile può avere infiniti valori, $n = \infty$)

Ho una quantità di dati SEMPRE inferiore ai 100 anni: non posso fare un'analisi statistica né lavorare sempre con $P = 0{,}99$ perché avrei dei valori sempre troppo alti.

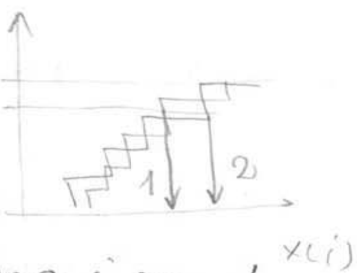
- VARIABILITÀ CAMPIONARIA



Ipotizzo di avere $n = 8$, $R = 1/100$ MA 2 anni dopo succede l'evento che sormonta l'opera.
Solo dopo scopro di avere 22 anni di dati ed aggiungo la seconda scala temporale.
Se ho 2 gruppi di dati, anche se con lo stesso numero, posso avere dei valori differenti.

La variabilità campionaria è intrinseca nel dato: anche quando diversi campioni vengono estratti dalla stessa popolazione, essi possono presentare distribuzioni di valori diversi.

Come tutelarci quando le $F(x)$ sono diverse?

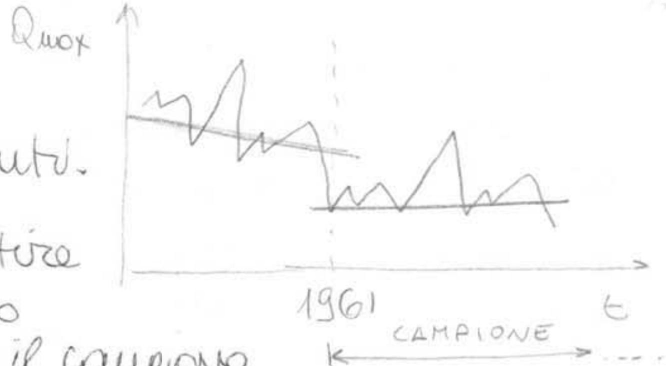


Anche se ho livelli di rischio elevati, il campione può produrre una stima molto distorta (se ragioniamo solo in modo empirico): devo ricercare la popolazione dietro i campioni.

Se ho che $R = 1/10$ ed $n = 22$ anni (esempio precedente), non basta scegliere il 2° valore MASSIMO, ma bisogna fare un'analisi più approfondita.

Bisogna collegare i punti solo quando sono continui nel tempo: se abbiamo delle carenze di dati, NON bisogna collegarli.
Possiamo trovare anche 1 singolo dato isolato: il problema del riempimento dei dati è che le serie possono essere brevi e discontinue; NON bisogna collegare i dati discontinui.

- STRANEZZE



La serie storica è una sequenza continua di punti.

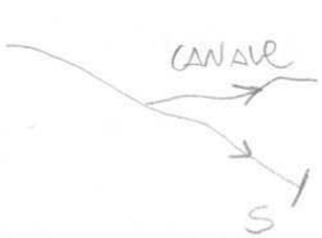
1) Questo caso ci fa insospettire perché i valori oscillano su 2 linee differenti: il campione NON È ESTRATTO DALLA STESSA POPOLAZIONE, ma da due.

Cos'è successo?
Sul corso d'acqua è stato fatto un invaso artificiale, a partire dal 1961: gli invasi producono una riduzione dei colmi di piena quindi dopo il 61' le piene sono state attenuate dalla presenza dell'invaso.

Siccome ci interessano le caratteristiche attuali, considero solo i dati post 1961.
Come faccio a capire se è funzione dell'invaso o della variabilità campionaria?
Il grafico ci fa venire il sospetto: bisogna verificare cos'è successo facendo un INCREMENTO DI INDAGINE.

ESEMPIO

In questo caso Q_{max} diminuisce vistosamente solo se il nostro corso d'acqua è piccolo.

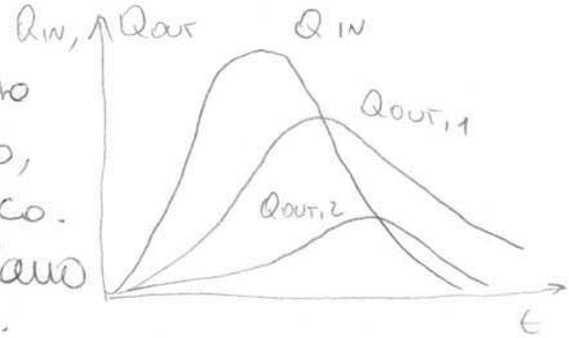


· EFFETTO DI LAMINAZIONE DELLA PIENA

Effetto per il quale, grazie alla disponibilità di volume (W_{disp}), il colmo di piena può essere attenuato.

W_{DISP} = VOLUME DISPONIBILE ALL'INVASO, la dimensione dipende se la piena trova l'invaso pieno / vuoto.

Se l'evento di piena trova l'invaso vuoto, il colmo di piena riempie l'invaso.



- $Q_{OUT,1}$ = volume disponibile basso
- $Q_{OUT,2}$ = volume disponibile alto, ho un basso livello idrico.

L'invaso può essere vuoto se riusciamo a prevedere l'evento di piena.

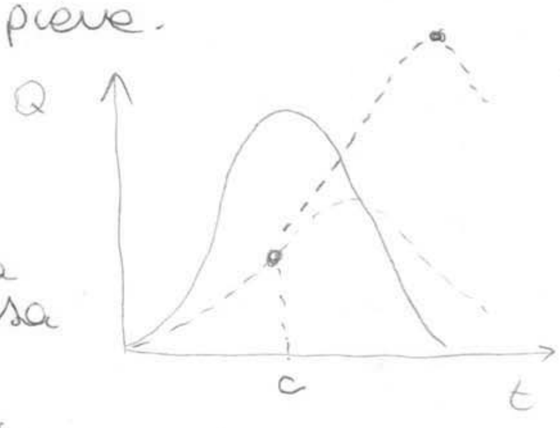
Nella realtà esiste sempre un volume disponibile d'invaso perché nelle dighe il livello d'invaso deve essere inferiore di almeno 1 m al LIVELLO DI MASSIMA REGOLAZIONE.

Il LIVELLO DI MASSIMO INVASO: livello nominale previsto dalla diga.

Bisogna garantire sempre un volume d'acqua disponibile.
$A = 1 km^2 = 10^6 m^2$, $h = 1 m \rightarrow 10^6 m^3$, valore significativo per contenere le piene.

· OSSERVAZIONE

Se succede un'errata valutazione sugli organi di scarico (C), aprendo la paratoia, l'acqua si riversa nel corso d'acqua.



Quando vedo che il livello d'acqua cresce velocemente, è sbagliato aprire gli scarichi perché avrei un passaggio di un grande volume d'acqua dall'invaso al corso d'acqua.

2- L'effetto climatico ha effetti INCONTESTABILI sulle temperature, mentre sulle altre variabili (come le piogge, abbiamo delle forti incertezze), perché NON ho né correlazioni, né una CHIARA EVIDENZA STATISTICA.

3- Nei bacini d'alta quota, l'effetto di un aumento di temperatura può essere molto significativo, per la presenza di neve.

linea a÷a = LINEA DELLO ZERO TERMICO

Sopra a÷a l'acqua si accumula e si può sciogliere in tempi diversi da quelli di piena: bisogna considerare lo SCIOGLIMENTO della neve, anche se, da solo, NON darà mai un picco di piena, perché è un valore piccolo MA CONTINUO.

Aumentando la T ho uno spostamento di a÷a, che produce un aumento dell'effetto di piena perché ho delle aree contribuenti più significative.
I bacini d'alta quota sono sensibili alle ΔT.

SOSPETTIAMO l'esistenza dei trend:

- Perché è un problema?
  E' come se ogni punto fosse estratto da una popolazione diversa: nel tempo cambiano le caratteristiche della popolazione quindi ogni anno cambia la popolazione.
- Quale dato prendo?
  NON posso prendere l'ULTIMO dato perché il valore continuerebbe a crescere nel tempo.

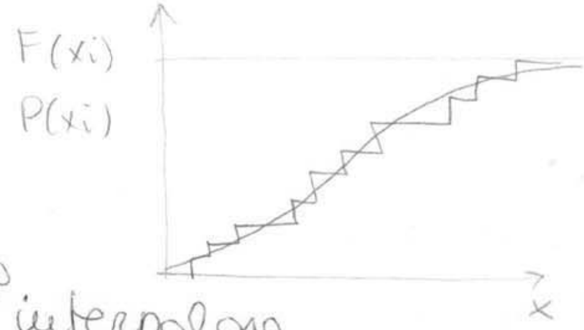
↳ METODI IN CONDIZIONI NON STAZIONARIE

$P(x,t)$, funzione del punto e del tempo.

B) SCELTA DEL MODELLO PROBABILISTICO

Scelta della distribuzione di probabilità che vogliamo rappresentare per simulare l'andamento del nostro campione.

$F(x_{(i)})$ funzione di frequenza empirica



Bisogna trovare un modello matematico che permetta di interpolare l'andamento del campione e di estrapolarlo nella parte di nostro interesse (coda destra della distribuzione) con probabilità 95÷99 %.

CARATTERISTICHE DELLA FUNZIONE $F(x)$

1) $$\boxed{\lim_{x \to x_{INF}} F(x) = 0}$$

Se $x \to x_{inf}$, valore minimo, $F(x) \to \emptyset$.

Questo vale perché $F(x)$ corrisponde ad una probabilità di non superamento.

- PORTATE $x_{INF} = 0$, non ha senso parlare di portate negative; x_{INF} ha INTERESSE IDROLOGICO
- TEMPERATURE $x_{INF} = -273\,°C$.

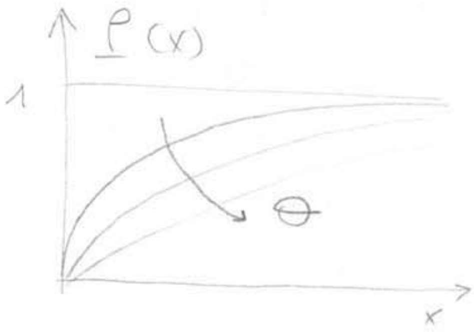
$F(x)$ deve trattare variabili positive.

2) $$\boxed{\lim_{x \to x_{sup}} F(x) = 1}$$

x_{sup} è il valore massimo per le variabili di nostro interesse, ovvero $x_{sup} \to +\infty$ perché tratteremo di variabili illimitate superiormente.
Non esiste un limite superiore alla portata.

Se $x_{(N)} = x_{MAX}$, $F(x_N) = 1$ CERTEZZA DI NON ESSERE SUPERATO

- DENSITÀ DI PROBABILITÀ DI X



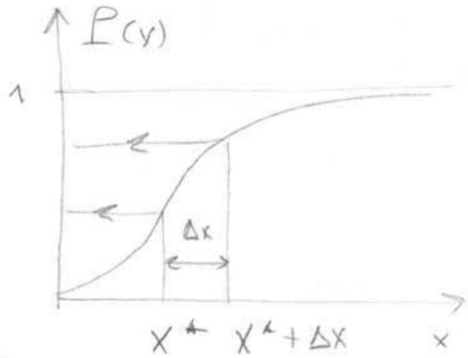
Al crescere di θ la curva si sposta verso il basso.

Il ruolo di θ è quello di ottimizzare l'adattamento del campione con la distribuzione.

DENSITÀ DI PROBABILITÀ

$$\boxed{\rho(x) = \frac{dF(x)}{dx}}$$

- Differenza tra $F(x)$ e $\rho(x)$?



Considero una generica distribuzione di probabilità: la $F(x)$ dice che è la probabilità che i valori non superino x^*.

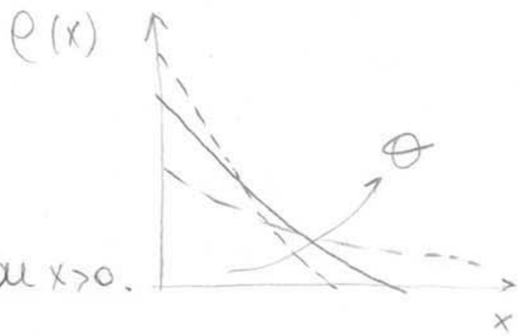
Qual è la probabilità di trovarsi nell'intorno di x^*?

$$\rho(x) = \frac{F(x^* + \Delta x) - F(x^*)}{\Delta x}$$

Perché divido per Δx? Perché altrimenti la $\rho(x)$ sarebbe funzione di Δx, invece è indipendente dall'ampiezza dell'intervallo.

Se facessi $\lim_{\Delta x \to 0} \rho(x)$ otterrei la definizione di derivata.

$\rho(x)$ rappresenta la probabilità di essere nell'intorno di un certo valore.



Se $x=0$ $\rho(x) = \frac{1}{\theta}$, poi decresce con $x>0$.

Le curve si ottengono con valori di θ differenti: aumentando θ ho una curva più ripida.

Fissato un livello di probabilità - $q = 0,99$.

$$q = P(x^*) = 1 - e^{-x^*/\theta} \rightsquigarrow x^* = -\theta \ln(1-q)$$

ricaviamo il valore di progetto per il prefissato livello di probabilità.
Però, per poter completare il progetto bisogna stimare il valore di θ.

Vediamo se la distribuzione esponenziale può funzionare per rappresentare i nostri dati.

c) VERIFICA PRELIMINARE TRAMITE CARTE PROBABILISTICHE

Bisogna fare una prima selezione dei modelli probabilistici, eliminando quelli che sono inadeguati per descrivere il campione.
Quest'operazione viene fatta PRIMA di determinare il valore dei parametri (θ_i) : non calcolo θ per tutti i modelli MA uso uno strumento grafico.

COSTRUZIONE CARTA PROBABILISTICA

1) ORDINARE il campione in senso crescente $x_i \to x_{(i)}$ dal valore più piccolo al più grande.
$x_{(i)}$ è lo i-esimo valore nel campione ordinato
2) ATTRIBUIRE una frequenza di non superamento (q_i) ai valori $x_{(i)}$

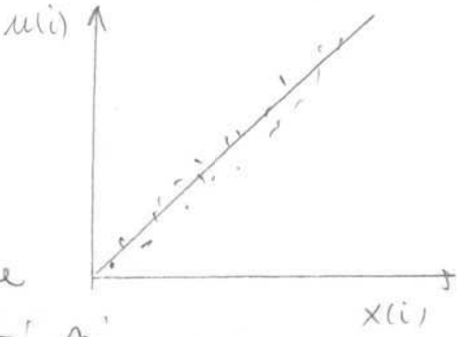
$$\boxed{q_i = F(x_{(i)}) = \frac{i}{N+1}}$$

ANALOGA ALLA FUNZIONE DI FREQUENZA EMPIRICA.

N : dimensione del campione

Al valore $x_{(N)}$ avrei che $q_i = 1$ MA io non voglio che $x_{(N)}$ sia superato: uso $N+1$

5) Rappresentazione grafica di $u(i)$ vs $x(i)$



Carta Probabilistica esponenziale.

Come ci aiuta?
Se effettivamente il campione è stato estratto dalla popolazione con distribuzione esponenziale, i punti si ALLINEANO sul diagramma: ragionevole allineamento dei punti sulla carta probabilistica.

Se ciò non avviene, scarto il modello probabilistico

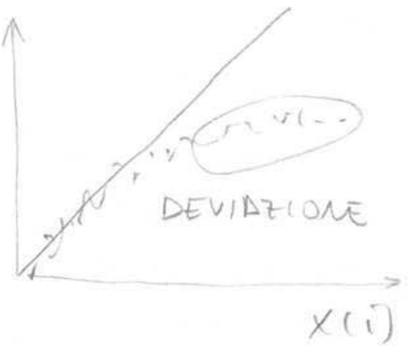
Perché i punti si allineano?

$u = \frac{x}{\theta}$ equazione di una retta passante per l'origine

Verificando l'allineamento si possono verificare tutti i passaggi fatti prima.

Quando uso $q(i)$ ho ipotizzato di essere nel mondo esponenziale: se trovo una relazione lineare tra $(u(i), x(i))$, significa che il modello esponenziale si presta bene al nostro campione.

ESEMPIO PUNTI NON ALLINEATI



Non posso rappresentare TUTTI I PUNTI. perché i punti non sono allineati perfettamente.

- ERRORI DA NON FARE

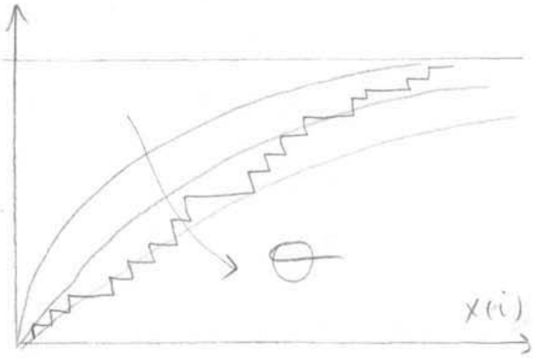
Non bisogna dimenticarsi gli steps e passare da

$u = \frac{x}{\theta}$ al valore $u = \frac{x_i}{\theta}$ con θ

parametro assunto.
Facendo così otterrei dei punti perfettamente allineati lungo una retta (IMPOSSIBILE).

D) STIMA DEI PARAMETRI DELLE DISTRIBUZIONI

$F(x_{(i)})$ FUNZIONE DI FREQUENZA EMPIRICA



Su $F(x_{(i)})$ aggiungiamo la curva $f(x)$ e ragioniamo nella distribuzione esponenziale.

Al variare di θ la curva si sposterà.

Il metodo di stima dei parametri è quello di cercare il migliore adattamento tra l'andamento dei punti di $P(x)$ e il nostro modello campionario $F(x)$: rendere più simile l'evidenza campionaria del campione con la popolazione.

Al posto del colpo ad occhio ....

METODI DI STIMA

1) METODO DEI MOMENTI
2) METODO DEGLI ELEMENTI

Il metodo della massima verosimiglianza non è un metodo efficiente in campo idrologico perché ho un numero limitato di dati.

1) METODO DEI MOMENTI

Bisogna eguagliare i momenti del campione con i momenti della distribuzione.

- MOMENTO CENTRALE DI ORDINE r - PER IL CAMPIONE -

$$\boxed{m_r = \frac{1}{n} \sum_{i=1}^{n} (x_i - \overline{x})^r}$$

$\overline{x} = \frac{1}{n} \sum_{i=1}^{n} x_i$ MEDIA CAMPIONARIA, è un momento di ordine 1 MA NON è centrale, bensì calcolato rispetto l'origine.

$\overline{x} = m_1'$

Nella realtà, m, M sono le stesse scritture, MA si riferiscono a 2 mondi diversi.

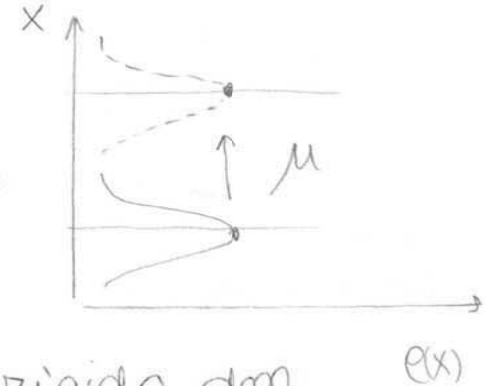
$$M_1' = \mu = \int_{ALL\,x} x\, \varrho(x)\, dx \quad \text{MEDIA DELLA DISTRIBUZIONE}$$

$$M_1 = \int_{ALL\,x} (x-\mu)\, \varrho(x)\, dx = \int_{ALL} x\, \varrho(x)\, dx - \mu \underbrace{\int_{ALL\,x} \varrho(x)\, dx}$$

$$= \mu - \mu \cdot 1 = \emptyset$$

dove $\int_{ALL\,x} \varrho(x)\, dx = 1$, area sottesa alla curva

I punti sono più densi nel valore centrale e poi la probabilità decresce: i valori più probabili sono quelli nel valore centrale mentre gli estremi convergono a zero.



Spostando μ, ho uno spostamento rigido della distribuzione senza cambiare la forma.

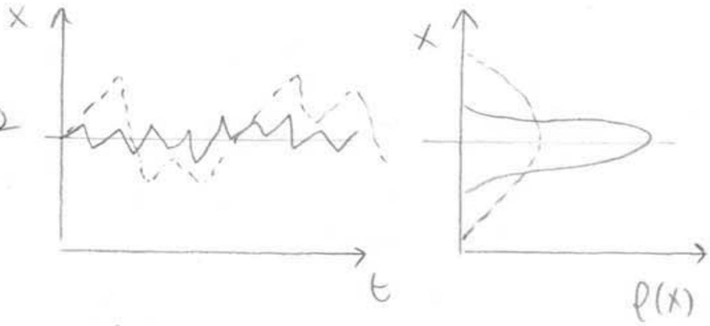
- MOMENTI DI ORDINE SUPERIORE

MOMENTO DI ORDINE 2

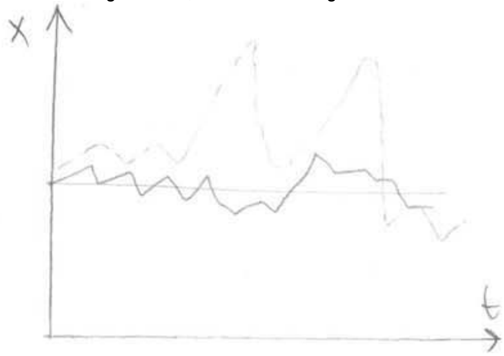
$$m_2 = \frac{1}{n} \sum_{i=1}^{n} (x_i - \bar{x})^2 = S^2 \quad \text{SCARTO QUADRATICO MEDIO}$$

$$M_2 = \int_{ALL\,x} (x-\mu)^2\, \varrho(x)\, dx = \sigma^2 \quad \text{VARIANZA}$$

Considero campione e distribuzione con la stessa media MA momenti di ordine 2 diversi.



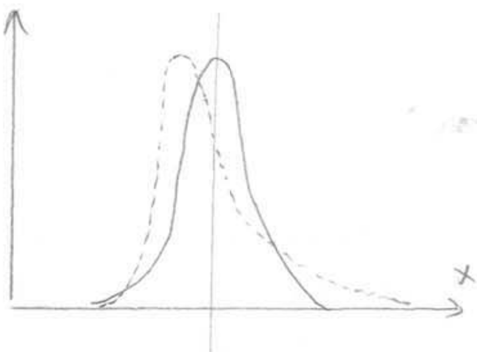
> la media è un indicatore di posizione.
> il momento del 2° ordine è un indicatore di DISPERSIONE intorno al valore centrale (INDICATORE DI SCALA) indica quanto i punti della distribuzione si discostano dalla media.



Considero media e dispersione costanti MA (M_3, m_3) variano.

— SERIE SIMMETRICA --- SERIE ASIMMETRICA

Nella serie asimmetrica ho una variazione dei dati (ad es. verso l'alto), che è diversa rispetto al valore centrale.



Ho una coda destra molto alta perché ho molti dati con $x > \bar{x}$.

— $M_3 = 0$ ho un asse di simmetria

--- $M_3 > 0$ non ho asse di simmetria.

I momenti di ordine 3 sono buoni per parlare di SIMMETRIA perché, mantenendo il segno, x^3 permette di esaltare le variazioni di valori presenti nella distribuzione.

Nell'ambito idrologico tratteremo distribuzioni asimmetriche

COEFFICIENTE DI ASIMMETRIA

$$CA = \frac{M_3}{\sigma^3} \qquad ca = \frac{m_3}{s^3}$$

Quando
- $CA = 0$ ho una distribuzione simmetrica
- $CA = 1$ ho una asimmetria marcata
- $CA > 2$ valori molto importanti.

ESEMPI

1) DISTRIBUZIONE ESPONENZIALE

$$F(x) = 1 - e^{-x/\theta}$$

Voglio stimare θ con il metodo dei momenti, uguagliando i momenti del campione con i momenti della distribuzione.

OSSERVAZIONI

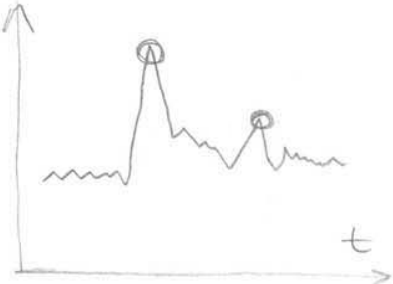
Nella realtà non è sempre così semplice.

1) Può succedere che gli integrali non siano risolvibili analiticamente: servono SOLUZIONI NUMERICHE.
Alcune dipendenze dei parametri dalla distribuzione NON sono risolvibili analiticamente.

2) La relazione può esplicitarsi analiticamente MA il sistema è NON RISOLUBILE (equazioni non lineari).
↳ sistemi di equazioni con soluzione numerica.

• PROBLEMA: OUTLIERS

I momenti di ordine elevato (sono $r \geq 2$) sono molto sensibili agli OUTLIERS: valori campionari che risultano molto discosti dagli altri valori presenti nella serie storica.



OUTLIERS: punti che differiscono dai valori tipici della serie temporale

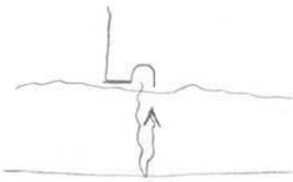
È frequente trovare gli outliers.

Bisogna capire se gli outliers sono il risultato di un processo fisico o derivano da qualche ERRORE

Questo ragionamento è complesso perché, se pensiamo agli eventi di piena, possono davvero esistere questi valori.

• MISURA DELLE PORTATE

1) MISURA TRAMITE IDROMETRO (h)



Storicamente sono delle aste graduate dove si legge il livello; oggi usano delle "padelle" che emettono onde e misurano il TIRANTE IDRAULICO.

Questa fase può essere soggetta ad errore, specialmente nei periodi di magra e a seguito di migrazione del corso d'acqua e erosione.

Nei piccoli corsi d'acqua ho una misura imprecisa.

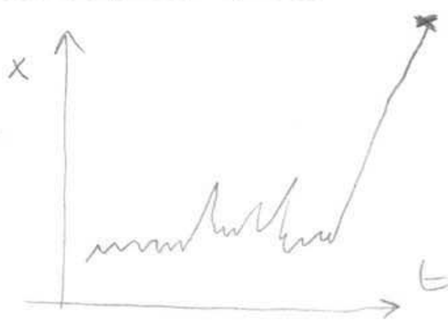
- OGGI: misura dei livelli in continuo e c'è la possibilità di fare una MODELLISTICA NUMERICA IDRAULICA per ricostruire la curva usando le equazioni dell'idraulica per andare a risolvere il problema di portate : livelli in maniera analitica..

Il problema sono i parametri di taratura, come la scabrezza che non è mai misurata MA va inserita come taratura del modello.
Il modello deve necessariamente avere delle misure DI CAMPO (pochi punti di taratura, 3/4).
I corsi d'acqua MIGRANO nel tempo e cambia anche la forma della sezione e la relazione tra velocità ed altezza: bisogna controllarla dopo gli eventi di piena.

L'incertezza del ±10% è intrinseca nei dati di portate e quando trattiamo le piene ricadiamo nella zona di ESTRAPOLAZIONE dove NON HO NESSUNA misura della velocità durante la piena.

Le misure di portata disponibili sono riferite ai periodi di magra, morbida (piena media) o di piena ordinaria dove non c'è pericolo.

L'incertezza sulle piene cresce tanto di più quanto l'evento diventa eccezionale (± 20 ÷ 30 %) e se l'evento è davvero eccezionale può anche saltare l'idrometro, cercheremo le tracce visibili del passaggio delle piene.



Bisogna controllare l'affidabilità del dato: STATISTICA AVANZATA.

METODO DEGLI L-MOMENTI (MOMENTI LINEARI)

Metodo che nasce dall'esigenza di rispondere al problema degli outliers e bisogna definire dei DESCRITTORI meno sensibili.

- DISTRIBUZIONE

$$\begin{cases} L_1 = B_0 \\ L_2 = 2B_1 - B_0 \\ L_3 = 6B_2 - 6B_1 + B_0 \end{cases}$$

L_i: combinazione lineare di altre grandezze

MOMENTI PESATI IN PROBABILITA' DI ORDINE r

$$\boxed{B_r = \int_{ALL\ x} x \, [F(x)]^r \, f(x) \, dx}$$

- OSSERVAZIONI

1) Nel confronto tra B_r, M_r noto che nell'elevamento ad r NON viene elevata la variabile (x) MA la densità di probabilità cumulata $F(x)$.

2) Le B_r sono lineari rispetto ad x e sono meno sensibili agli outliers; inoltre vengono usati gli L-momenti perché sono di facile interpretazione:
L_1 = MEDIA $\quad L_2$ = DISPERSIONE $\quad L_3$ = ASIMMETRIA

$$B_0 = \int_{ALL\ x} x \, f(x) \, dx = \mu$$

$L_1 = B_0 = \mu$ la media non è elevata ad r e c'è una relazione lineare tra μ ed x.

$B_1 = \int_{ALL\ x} x \, F(x) \, f(x) \, dx$ linearità con x.

Perché combino gli B_r per ottenere L_r?
Gli L-momenti sono di facile interpretazione ed hanno l'interpretazione analoga dei momenti classici; inoltre usiamo gli L-momenti anche perché esiste il legame lineare tra B_r ed x.

Due diverse distribuzioni che hanno valori diversi di L_1, L_2 hanno una diversa dispersione.
$L_3 > 0$ è l'antisimmetria, mentre $B_3 > 0$ non ha alcun significato.

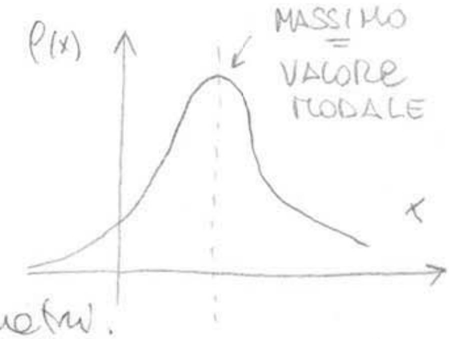
- OSSERVAZIONE

1) Se la serie è molto lunga, i risultati ottenuti con diversi stimatori tendono a convergere verso lo stesso valore. Con solo 30 anni di dati ho gli stimatori discosti.
2) Perché usiamo anche la stima dei momenti e non solo quella degli L-momenti?

- Motivo storico: bisogna confrontarsi con le altre persone ed avere un riferimento comune;
- INCERTEZZE: per controllare le incertezze e fare i calcoli con diverse strategie è quello di confrontare i dati e vedere di quanto si discostano i risultati.

DISTRIBUZIONE GAUSSIANA - NORMALE -

$$\boxed{p(x) = \frac{1}{\sqrt{2\pi}\,\theta_2} e^{-\frac{1}{2}\left(\frac{x-\theta_1}{\theta_2}\right)^2}}$$



la distribuzione normale ha 2 parametri.
la distribuzione normale è poco usata per gli eventi estremi, perché:

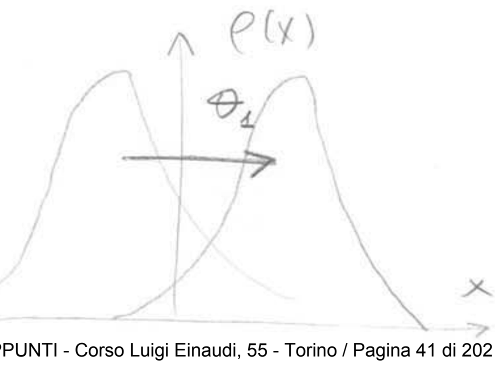
1 - $x \in]-\infty, +\infty[$, ricopre tutti i valori di x.
2 - $CA = 0$ coefficiente di asimmetria NULLO, perché la distribuzione è simmetrica.

la Gaussiana non si adatta bene agli eventi estremi perché tratta anche le variabili negative, mentre la portata ha solo valori positivi.
Usare una distribuzione simmetrica per descrivere eventi estremi (asimmetrici) non va molto bene.
la Gaussiana servirà per descrivere la lognormale

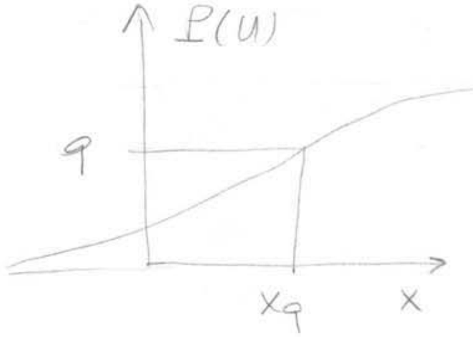
RUOLO DEI PARAMETRI

θ_1 = PARAMETRO DI POSIZIONE

Modificare θ_1 significa spostare rigidamente la distribuzione



Graficamente, imponendo q ottengo $x(q)$.



$$F(u) = \Phi(u)$$

$$u_q = \Phi^{-1}(q) = \text{INV.NORM.ST}(q)$$

5) PLOTTARE $u(i)$ VS $x(i)$

Dopo questo, cosa possiamo fare?
Bisogna definire un livello di rischio ed estrarre il quantile $x_q = \theta_1 + u_q \theta_2$, formula presa dalla relazione della variabile ridotta.

• CONFRONTO CON LA DISTR. GUMBEL

Anche la Gumbel ha un dominio

$$x \in]-\infty ; +\infty[$$

MA ha anche un ASIMMETRIA : $CA \simeq 1,14$
la distribuzione Gumbel è asimmetrica e permette di avere la CODA DI DESTRA molto più sviluppata della coda di sinistra : la parte con $x < 0$ è TRASCURABILE.

DISTRIBUZIONE LOGNORMALE

La distribuzione lognormale è la distribuzione di una variabile casuale il cui logaritmo della variabile ha una distribuzione normale:

$$y = \ln(x)$$

y ha una distribuzione normale
x variabile casuale

Come descrivere la lognormale?
Ragiono sulle DISTRIBUZIONI DERIVATE (VEDI DISPENSE)

$$\begin{cases} F(x) = F(u) \\ p(x) = p(y) \dfrac{dy}{dx} \end{cases}$$

STIMA DEI PARAMETRI DELLA LOGNORMALE

$\theta_1 = \mu(\ln(x))$ MEDIA DEI LOGARITMI DEI DATI

$\theta_2^2 = \sigma^2(\ln(x))$ VARIANZA DEI LOGARITMI DEI DATI

Il quantile $\ln(x_q) = \theta_1 + \theta_2 u_q$

$$\boxed{x_q = e^{(\theta_1 + \theta_2 u_q)}}$$

Fare il logaritmo dei dati permette di trattare la distribuzione lognormale come una gaussiana che serve per svolgere l'analisi statistica degli eventi estremi.

GENERALIZED EXTREME VALUE (GEV)

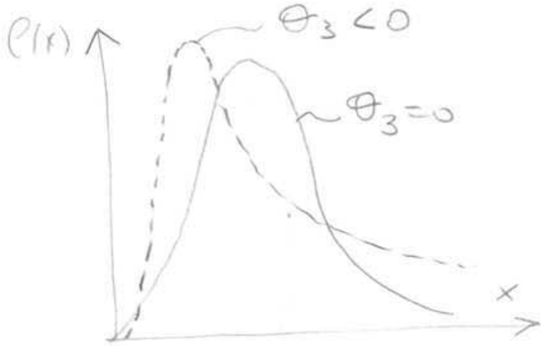
Distribuzione con 3 parametri.

$$\boxed{f(x) = \exp\left[-\left(1 - \frac{\theta_3}{\theta_2}(x - \theta_1)\right)^{\frac{1}{\theta_3}}\right]}$$

θ_1 = PARAMETRO DI POSIZIONE

θ_2 = PARAMETRO DI SCALA

θ_3 = PARAMETRO DI FORMA DELLA DISTRIBUZIONE



Per $\theta_3 = 0$, la GEV corrisponde alla Gumbel:

— $\theta_3 = 0$ CA = 1,14

---. $\theta_3 < 0$ CA > 1,14

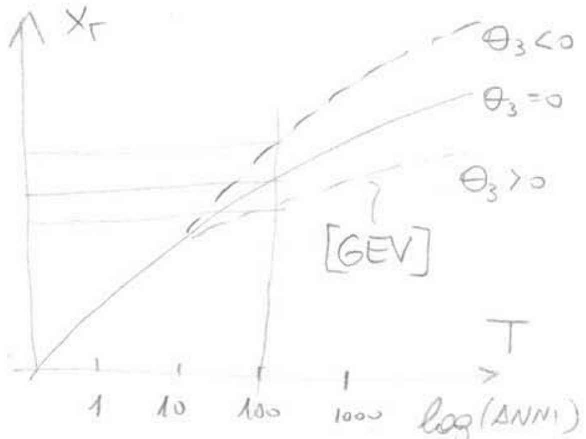
Il caso di $\theta_3 < 0$ è quello di nostro interesse.

Cosa succede se $\theta_3 < 0$?

- $x \in \left[\theta_1 + \frac{\theta_2}{\theta_3} ; +\infty\right[$

Il dominio è vincolato al limite inferiore, definito dai parametri θ.

ESEMPIO D. ESPONENZIALE



x_T : valore di progetto corrispondente ad un determinato T.

$$1 - \frac{1}{T} = 1 - e^{-x_T/\theta}$$

$$x_T = -\theta \ln\left(\frac{1}{T}\right) = \theta \ln(T)$$

Diagramma che permette di vedere bene cosa accade alla coda destra delle curve delle distribuzioni.

θ_3 ha una rilevanza importante nel progetto perché i valori tendono a divergere per piccole variazioni di θ_3.

La GEV ha una maggiore flessibilità rispetto alla GUMBEL perché permette di simulare meglio delle forme di dati più complesse.

GEV → + FLESSIBILE, si adatta meglio
→ − ROBUSTA, se ho pochi dati, stimo un valore di θ_3 MOLTO SBAGLIATO
la GEV si usa quando ho molti dati.

GUMBEL → + RIGIDA, + ROBUSTA } se ho pochi dati non sbaglio di molto.

- RUOLO DEL NUMERO DEI PARAMETRI DI UNA DISTRIBUZIONE

All'aumentare del numero dei parametri di un modello (θ) aumenta la capacità di adattamento del modello matematico ai dati ma diminuisce la sua robustezza.

ESEMPIO

(X, Z) variabili casuali che vogliamo mettere in relazione

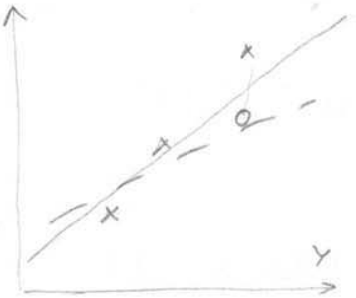
© Proprietà riservata dell'autore - Digitalizzazione e distribuzione a cura del CENTRO APPUNTI - Corso Luigi Einaudi, 55 - Torino / Pagina 47 di 202

la zona di estrapolazione è la zona nella quale non abbiamo osservazioni empiriche nel modello: in funzione allo spostamento di un punto avrò valori MOLTO discosti quando voglio valutare la zona di estrapolazione.
Il modello non è robusto quando ha MOLTA incertezza nelle zone di estrapolazione.

$r=0$ NON È ATTENDIBILE, si adatta poco bene ai dati

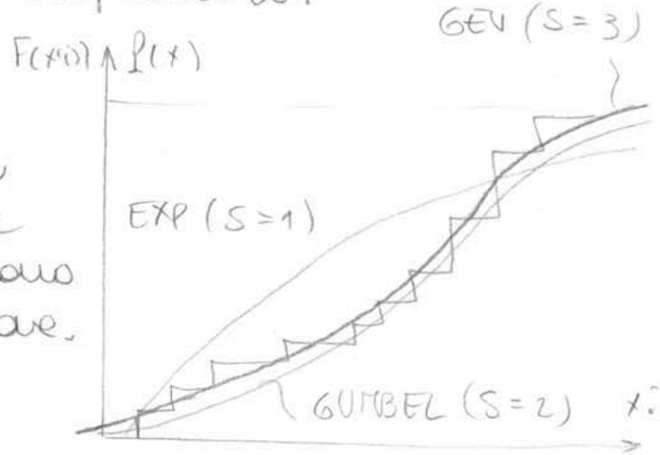
$r=2$ NON È ATTENDIBILE perché si adatta troppo bene ai dati, pericolo nelle zone di estrapolazione.

$r=1$ sarebbe il modello che si adatta meglio alle variazioni dei dati campionari.



Tornando a noi...

Quando trattiamo le portate, i nostri valori (funzione dei tempi di ritorno) ricadono nella zona di estrapolazione.



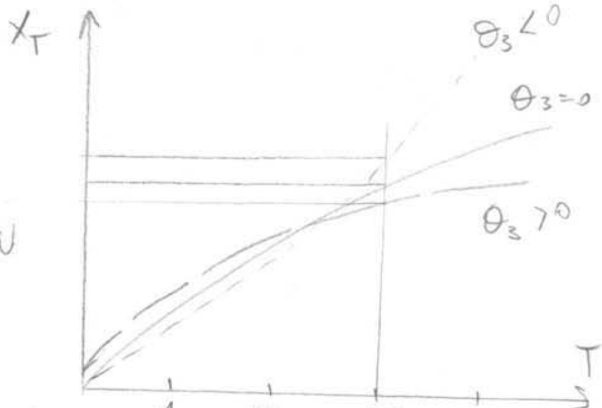
GUMB.: può avere lo stesso problema di $r=2$ in zone di estrapolazione

EXP: NON conosce bene i punti

GEV: ha un adattamento migliore della Gumbel MA ha Θ_3: se Θ_3 varia di poco, produce delle variazioni notevoli nei valori di progetto.

DISTRIBUZIONE GEV

Ho delle variazioni PESANTI dei valori di progetto SE scelgo la distribuzione GEV o la Gumbel ($\Theta_3=0$).



È sufficiente avere $n = 10/15$ anni di dati per usare la distribuzione Gumbel o exp.

E) VERIFICA TRAMITE TEST DI ADATTAMENTO:

Test statistici atti a verificare l'ipotesi statistica che il campione sia stato estratto da una certa distribuzione di probabilità nota.

TEST STATISTICI

1) TEST A 2 CODE

H_0 = IPOTESI STATISTICA ipotesi che un campione disponibile sia stato estratto da una distribuzione simmetrica (noi decidiamo l'ipotesi da verificare).

H_0 è l'ipotesi statistica soggetta a test: noi decidiamo quale ipotesi H_0 deve soddisfare.

Stimiamo il coefficiente di asimmetria del campione.

VARIABILE TEST

Variabile statistica che costruisco per valutare l'ipotesi: in questo esempio userò C.A.

$$CA^* = \frac{m^3}{S^3} = 0{,}714$$

È L'UNICO CHE POSSO USARE? $\frac{L_3}{L_2}$

L'ipotesi di H_0 è verificata?

CA^* da solo non basta per capire se il campione sia stato estratto da una distribuzione simmetrica.

Quando parliamo di variabili continue, anche se il campione è stato estratto da una distribuzione simmetrica, NON otteniamo MAI $CA = 0$, perché la variabilità campionaria è intrinseca nel campione.

CA^* = rappresenta che è probabile che il campione sia estratto da una distribuzione simmetrica.

CA' = è abbastanza improbabile estrarre da una distribuzione simmetrica un campione che mi fornisca proprio questo CA'.

• <u>LIVELLO DI SIGNIFICATIVITÀ (α)</u>

| Probabilità di compiere un errore di TIPO I.

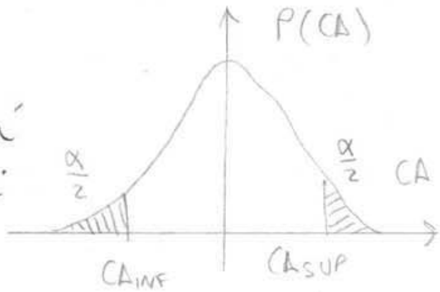
• <u>ERRORE DI TIPO I</u>

| Rigettare l'ipotesi Ho quando essa è vera.

Torniamo alla curva ($P(CA)$, CA) ed usiamo $\alpha = 0,10$.

CA_{SUP} = limite superiore di accettabilità

CA_{INF} = limite inferiore di accettabilità



Questi sono i limiti di accettabilità del test, ovvero è il limite che la probabilità ha di ricadere in $CA > CA_{SUP}$ o $CA < CA_{INF}$, limite pari ad $\frac{\alpha}{2}$.

Se $CA_{INF} < CA^* < CA_{SUP}$, il test è passato e l'ipotesi Ho è VERA, altrimenti l'ipotesi Ho è falsa.

$$P_{CA}(CA_{SUP}) = 1 - \frac{\alpha}{2} \qquad P_{CA}(CA_{INF}) = \frac{\alpha}{2}$$

$P_{CA}(CA)$ non è la probabilità della variabile X ma è la distribuzione di variabilità della CA, variabile test.

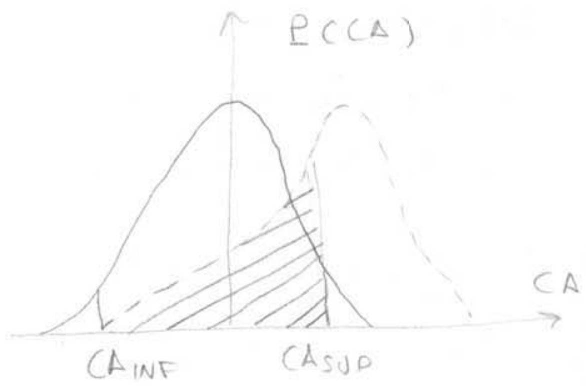
$CA_{SUP} = P^{-1}\left(1 - \frac{\alpha}{2}\right)$ quantile con probabilità $1 - \frac{\alpha}{2}$

$CA_{INF} = P^{-1}\left(\frac{\alpha}{2}\right)$ quantile con probabilità $\frac{\alpha}{2}$

| Ho è verificato se $CA_{INF} < CA^* < CA_{SUP}$.
| Ho è rigettata se $CA^* < CA_{INF}$ o $CA^* > CA_{SUP}$

Cosa serve H_1?

Se scelgo α troppo piccolo:
- diminuisce la probabilità di fare errori di tipo I;
- aumenta la probabilità di fare errori di tipo II.



- ERRORE DI TIPO II

Accettare Ho quando essa è FALSA.

L'area sottesa alla curva H_1, compresa tra i limiti (CA_{INF}, CA_{SUP}) corrisponde alla probabilità di fare errori di tipo II.

Quando $CA_{INF} < CA^* < CA_{SUP}$, potrei accettare l'ipotesi Ho quando questa è FALSA: non posso scendere a valori di α troppo bassi.

β = probabilità di fare errori di tipo II, area sottesa.

Se diminuisce la probabilità di fare errori di tipo I (α), aumenta la probabilità di fare errori di tipo II (β): chi effettua il test deve scegliere un valore di α tale da minimizzare gli errori di tipo I, II.

In generale: $0,01 < \alpha < 0,1$

$\alpha = 0,01 = 1\%$ test poco restrittivo, limiti ampi,
$\alpha = 0,1 = 10\%$ test più restrittivo

RIASSUNTO

1) Noto Ho, scegliere quale variabile usare come test (es. CA);
2) Scegliere α;
3) Determinare i limiti di accettabilità del test: i quantili $\left(\frac{\alpha}{2}\right)$, $\left(1-\frac{\alpha}{2}\right)$ e conoscere P^{-1}.

Bisogna suddividere il campione in classi, dove ogni classe conterrà un certo numero di elementi;

ESEMPIO

$n = 74 \qquad k = 2 \cdot 74^{0,4} \cong 11,2 = 11$ classi

$n\,P_i$ = NUMERO DI ELEMENTI ATTESO NELLA CLASSE i-ESIMA QUANDO VALE L'IPOTESI H_0.

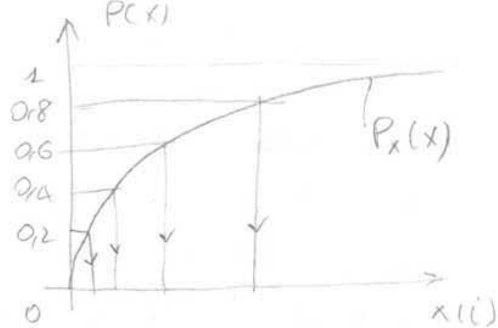
Il Test di Pearson confronta 2 grandezze:

- n_i per il campione
- $n\,P_i$ per la popolazione

- Se X^2 è piccolo, significa che campione e popolazione si assomigliano ed il test è passato;
- Se X^2 è grande, significa che campione e popolazione sono distanti, il test non è passato.

X^2 è un test ad una sola coda perché $X^2 > 0$.

ESEMPIO



H_0 è una distribuzione generatrice esponenziale.

Applico la suddivisione in classi.

$P_i = \frac{1}{k}$ CLASSI EQUIPROBABILI, la probabilità di ricadere in una classe è costante.

$P_x(x) = 1 - e^{-x/\hat{\theta}}$ distribuzione di probabilità della variabile casuale con $\hat{\theta}$ stimato.

$k = 5 \qquad P_i = \frac{1}{5} = 0,2 \quad , \; 5$ classi.

le classi sono equispaziate lungo $P(x_i)$ ma non sono equispaziate lungo $x(i)$.

- $X^2 < X^2_{LIM}$ Ho è ACCETTATA
- $X^2 > X^2_{LIM}$ Ho è RIGETTATA

- OSSERVAZIONI

1) Cosa cambia se considero distribuzioni diverse?
Cambiano le definizioni dei limiti delle classi e cambiano i valori di n_i (numero di valori campionari nella classe i-esima). Come mai?
↳ Perché i limiti delle classi sono funzione dei quantili delle diverse distribuzioni.

Inoltre il metodo cambia se uso Ho con Momenti o L-Momenti, è possibile che il test accetti solo 1 dei 2 metodi.
→ otterremo 2 curve leggermente diverse, funzione della differenza dei parametri

2) $P_{X^2}(X^2)$ DIPENDE dal NUMERO DEI PARAMETRI (S)

S=1 EXP S=2 GUMBEL S=3 GEV ...

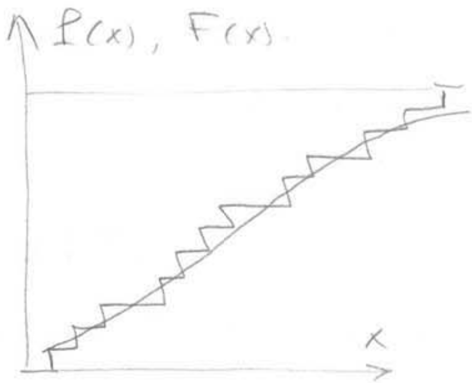
— S=3 - - - S=1

$X^2_{LIM,S=1} > X^2_{LIM,S=3}$ la curva si sposta

Con meno parametri ho il limite molto più a destra.

Quando ho più parametri faccio un test molto più restrittivo: cerco di essere più "severo" nelle distribuzioni che si adattano meglio ai valori.

$f(x) - F(x)$ = indicazione della distanza tra le curve.



Se ho distanze piccole, avrò un miglior adattamento.

In questo test non ho le CLASSI.

$$W = \frac{1}{f(x)(1-f(x))}$$ FATTORE DI PESO

Se $f(x) = 0{,}5 \rightarrow W = 4$, parte centrale della curva

Se $f(x) = 0{,}99 \rightarrow W \simeq 100$, parte della coda destra.

Metto un peso maggiore nella coda di destra rispetto alla parte centrale della distribuzione.

A^2 E' SENSIBILE alla differenza $(f(x) - F(x))$, specialmente nella coda di destra.

$$\boxed{A^2 = -n - \frac{1}{n} \sum_{i=1}^{n} \left[(2i-1) \ln(f(x_{(i)})) + (2n+1-2i) \cdot \ln(1 - f(x_{(i)})) \right]}$$

Equazione equivalente all'integrale dove dentro c'è $P(x_{(i)})$ = distribuzione di probabilità della variabile sottoposta a test (da calcolare con i campioni ordinati).

IMPORTANTE A^2 E' UN NUMERO

Se $A^2 < A^2_{LIM}$, Ho è accettata.

Posso eseguire un cambio di variabile:

$$W = f(A^2, \xi_p, \beta_p, \gamma_p)$$ VEDI DISPENSE.

A - UTILIZZARE IL VALORE PIÙ ALTO

È vero che scegliere il valore più alto significa mettersi in sicurezza MA significa anche DISTORCERE il valore di progetto perché prendiamo valori con un maggiore tempo di ritorno rispetto al valore fissato.

Se ci spostiamo su un tempo di ritorno più ampio ($T = 100 \rightarrow T = 140$), siamo più sicuri MA: aumentano i costi dell'opera e DIMINUISCONO i danni attesi, MA questa NON È la RICHIESTA del committente.

Andare verso fattori di sicurezza maggiori significa aumentare i costi dell'opera.

B - MEDIARE LE STIME OTTENUTE

$$X^* = \frac{x_1^* + x_2^* + \ldots + x_n^*}{n}$$ perché ho più valori x_n^*?

Questo perché ho un incertezza residua che non riesco ad eliminare ed è intrinseca nei dati.
Non sempre abbiamo l'effettiva convergenza tra campione e popolazione di provenienza: il campione può essere stato estratto da 1 delle diverse distribuzioni; devo fare la MEDIA tra le stime ottenute.

C'è un fattore di soggettività in queste analisi, in ogni fase di scelta iniziale: se 2 persone fanno le stesse analisi sui dati, ottengono valori differenti; questo perché c'è un incertezza intrinseca nelle analisi considerate, e rischiamo di convergere verso valori molto differenti.
Come fare per limitare le incertezze?
Bisogna cercare di essere i più generali possibili usando un maggior numero di metodi (bisogna sempre fare l'analisi costi/benefici).

Se aumentano le incertezze, bisogna aumentare il numero delle distribuzioni e fare TEST PIÙ AVANZATI.

Q é una variabile cumulata che risente dell'accumulo del bacino idrografico.

Faccio l'analisi statistica in D.

Nel passaggio da D a S devo considerare che ho un aumento del bacino idrografico.

A_D = bacino idrografico che concerne alla sezione D

$\frac{A_S}{A_D} > 1$ perché A_S ha un bacino idrografico > di A_D.

Area da aggiungere al bacino A_D per ottenere A_S.

Stesso discorso se D é a valle di S.

- LIMITI DI APPLICABILITA'

1) $\boxed{0,5 < \frac{A_S}{A_D} < 2}$ VALORI INVENTATI

Posso allontanarmi finché 1 area di 1 bacino non é il doppio rispetto l'area dell'altro bacino.

$A_S/A_d = 1,1$ posso stare tranquillo

$A_S/A_d = 1,9$ é meglio fare l'analisi AFFLUSSI/DEFLUSSI

- EFFETTO DEL NUMERO DEI DATI IN D

- Se in D ho 80 anni di dati, l'inferenza statistica funziona bene, della stima mi fido e posso propagare la stima in S.
- Se in D ho 15 anni di dati, bisogna stare attenti e cercare di avere sezioni vicine

2) $\boxed{L_{SD} < 10 \text{ km}}$ DISTANZA TRA LE SEZIONI

Perché devo rispettare le 2 condizioni?

Perché S, D devono essere vicine?

Non possiamo solo basarci sul valore di β_1 per dire che/se il trend sia significativo: bisogna capire quanto β_1 può essere diverso da zero.

2) TEST DEL T DI STUDENT

IPOTESI DA VERIFICARE $\boxed{H_0 : \beta_1 = 0}$

VARIABILE USATA PER EFFETTUARE IL TEST:

$$\boxed{t = \frac{\beta_1 \sqrt{n-2}}{\sqrt{\dfrac{\sum_{i=1}^{n} (x_i - \bar{x})^2}{\sum_{i=1}^{n} (t_i - \bar{t})^2}}}}$$

x_i = PORTATA
t_i = TEMPO

$\bar{t}$ è il tempo medio, infatti tratto il tempo come VARIABILE STATISTICA v.

Tanto più $\beta_1 \neq 0$, tanto più H_0 è da rigettare.

(A) Se ho piccole oscillazioni, allora anche un piccolo valore di inclinazione nei minimi quadrati può essere significativo.

(B) Se ho grandi oscillazioni, una piccola inclinazione dei minimi quadrati può anche non essere significativa (posso non avere trend).

β_1 non è sufficiente per dire se le fluttuazioni sono significative o meno.

Il Test della T di Student è un test a 2 code perché posso valutare trend positivi e negativi.

PRECIPITAZIONI

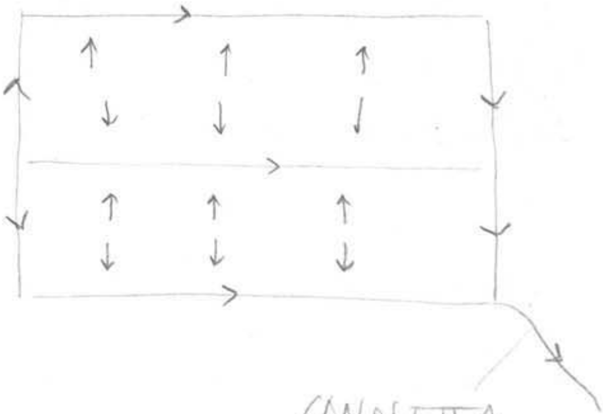
Capiamo come usare le informazioni pluviometriche per il progetto delle opere idrauliche.

PRECIPITAZIONI DI PROGETTO (VERIFICA)

Fonte di dati da usare per il dimensionamento delle opere idrauliche.

ESEMPIO: PARCHEGGIO

Dreno l'acqua in una canaletta.
Devo determinare le dimens. della canaletta per lo scolo delle acque di pioggia.



Q è la variabile di progetto.

Parliamo di tempo di ritorno: Q^* con $T = 5$ anni.
Perché T è così basso?
Questi sistemi vanno in crisi perché non viene fatta una buona manutenzione.

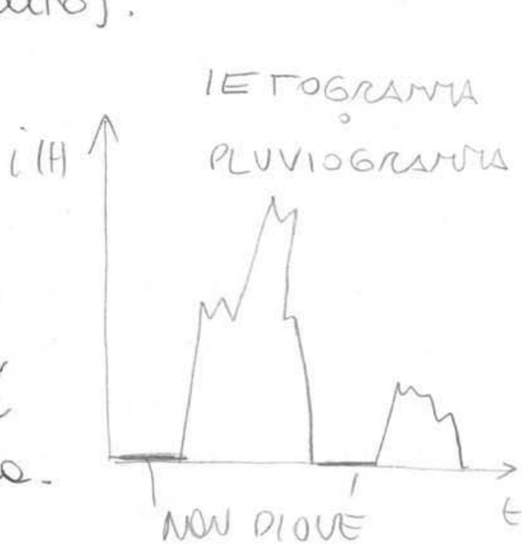
IMPORTANTE: NON HO I VALORI DI Q PERCHÉ NESSUNO MISURA LE PORTATE IN UN PARCHEGGIO.

Bisogna riferirsi alle precipitazioni per ricostruire le portate (variabile di riferimento).

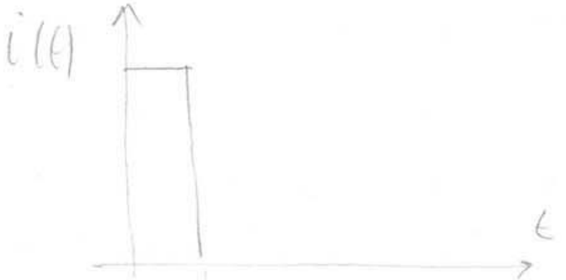
PLUVIOMETRO vicino al parcheggio: in continuo misuro il dato di precipitazione.

$i(t)$ = INTENSITÀ DI PRECIPITAZIONE

$[i(t)] = [mm/h]$ è la velocità, mm d'acqua caduti all'ora.

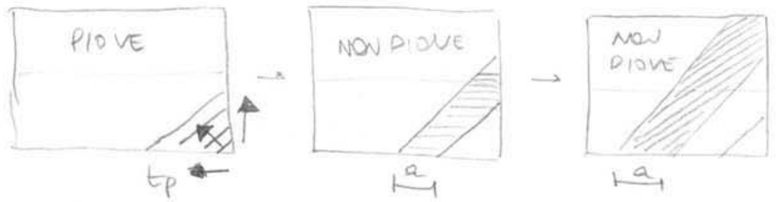


AREE CONTRIBUENTI



EVENTO 1

Evento molto intenso ma di breve durata: le aree contribuenti crescono in maniera lineare con il procedere del tempo.

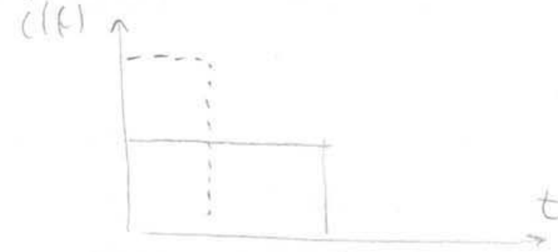


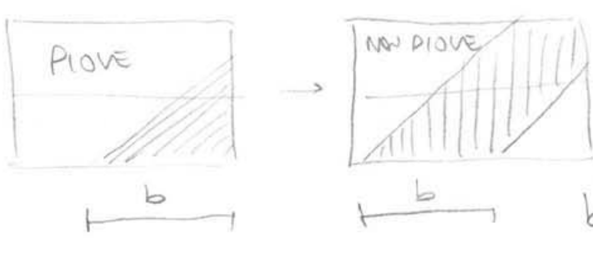
t_p = tempo di pioggia: per $t > t_p$ non piove più

Dopo t_p come si comportano le aree contribuenti?
la parte più vicina smette di contribuire e l'area contribuente si sposta/allontana dalla condotta: considero aree che non contengano più le zone vicine. Quando l'area contribuente esce dal bacino idrografico, le portate decrescono.

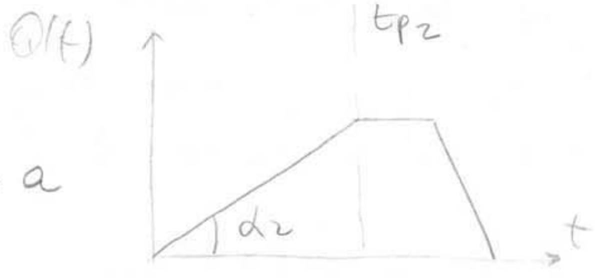
EVENTO 2

Evento meno intenso ma di maggiore durata





$b > a$



$t_{p_2} > t_{p_1}$; $b > a$

Ho delle aree contribuenti molto più grandi. le portate crescono con una minore pendenza ($\alpha_2 < \alpha_1$) perché la pendenza di crescita delle portate è funzione dell'intensità delle piogge ma $t_{p2} > t_{p1}$.

BACINO PICCOLO: piogge intense, tempi brevi
BACINO GRANDE: Bisogna guardare bene il corso di INTENSITA' PICCOLE CON PIOGGE PROLUNGATE (durante i mesi autunnali ho le piogge di lunga durata).

Se le aree contribuenti interessano un intero bacino idrografico avrò eventi dannosi.
Un temporale estivo (istantaneo) non interferisce dannosamente su un grande bacino idrografico (Po) MA manda in crisi i piccoli bacini.

DURATA CRITICA

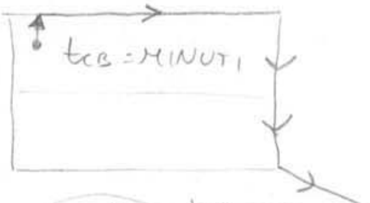
la durata critica è la durata della precipitazione che produce il maggior impatto sul bacino idrografico considerato

$$\boxed{d_{CR} = t_{CB}}$$

t_{CB} = TEMPO DI CORRIVAZIONE DEL BACINO: tempo che impiega una goccia d'acqua che cade nel punto IDRAULICAMENTE più lontano per raggiungere la sezione di chiusura.

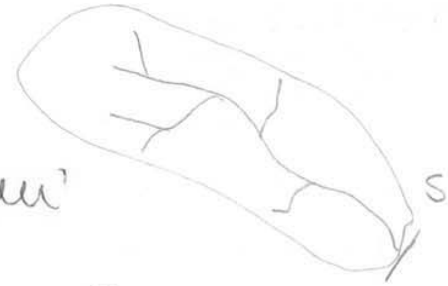
Qual è il punto idraulicamente più lontano?

Il punto idraulicamente più lontano NON corrisponde al punto geograficamente più distante: è il punto a cui corrisponde un maggior tempo di percorrenza della goccia d'acqua per raggiungere la sezione di chiusura.



t_C = TEMPO DI CORRIVAZIONE (valido per ogni punto) è il tempo che impiega la goccia d'acqua per raggiungere, dal punto di caduta della goccia, la sezione di chiusura.

Il t_C vale per ogni punto nel bacino idrografico.



Quindi nelle precipitazioni devo considerare la durata critica, valore funzione delle dimensioni del bacino idrografico.
Come posso usare questa informazione?

- IPOTESI: Non sono disponibili le misure DIRETTE delle portate in S, neanche nelle sue vicinanze.
  Voglio progettare l'opera con $T = 20$ anni.

DIMENSIONAMENTO OPERA IN S

1) DEFINIZIONE DELLA DURATA CRITICA

Sono presenti solo i dati delle precipitazioni:

$$d_c = t_{cB}$$

Perché assumere questa relazione? VEDI DOPO...

t_{cB} è una grandezza che può essere stimata INDIPENDENTEMENTE dalle precipitazioni: è una caratteristica del bacino idrografico e non è funzione del tipo di pioggia.

t_{cB} viene calcolato con formule empiriche: in generale
$t_{cB} \propto \sqrt{A_b}$ con A_b area del bacino.

A parità di velocità di scorrimento, t_{cB} è funzione della distanza percorsa, quindi è proporzionale all'area del bacino idrografico.

2) l'utilizzo delle precipitazioni è MOLTO USATA nel progetto, anche quando le opere sono poco importanti.

Se facciamo un lavoro PICCOLO non dobbiamo usare analisi molto costose: serviamo dei metodi semplici e di facile applicabilità.

Come possiamo affrontare il PUNTO 2?
Devo fare riferimento ai dati PRE-INTRODUZIONE della raccolta automatica / continua.

RACCOLTA E PUBBLICAZIONE DATI

- SERVIZIO IDROGRAFICO NAZIONALE (fino 1/2 anni '90)
- ARPA, AGENZIA NAZIONALE PER L'AMBIENTE (dopo 1/2 anni '90)

Negli anni 90 ho pochi dati: il SIN non funzionava male MA si hanno avuti problemi nel passaggio.

Il SIN come pubblicava i dati?
L'ARPA attualmente pubblica i dati online, mentre il SIN, per le piogge, pubblicava i dati negli ANNUALI IDROLOGICI.
Cosa troviamo in queste tabelle?
Troviamo i MASSIMI ANNUI DI PRECIPITAZIONE per le durate pari a 1, 3, 6, 12, 24 h.

| L'informazione dei dati PRE-ARPA non è continua MA SOLO funzione di queste 5 classi.

Il SIN forniva questa tabella:

INTENSITA' MEDIA in mm/h (I MASSIMI DELLE MEDIE)

	1h	3h	6h	12h	24h
1935	70	50	40	25	15
1936	120	55	38	22	13

Dati letti e collegati MANUALMENTE (prima dell'ARPA)

3) Il modo con cui decrescono i dati ci dice che tipo di evento è: nel 1936, dopo un intenso evento di 120 mm/h, ho avuto una forte attenuazione.

4) I numeri devono essere correlati.

$1\,h = 120\ mm/h$, $3\,h = 55\ mm/h$ $\qquad 3h > \frac{1}{3} \cdot 120 = 40$

E' importante verificare di avere una relazione tra i dati successivi!

$3h = 55\ mm/h \qquad 6h = 40\ mm/h > \frac{55}{2}$ OK

$6h = 20\ mm/h < \frac{55}{2}$ ERRORE!

Il nostro metodo BASE deve fare riferimento all'informazione contenuta in queste 5 durate.

2.A) CASO FORTUNATO

d_C COINCIDE con una delle 5 durate disponibili

es. $d_C = 2{,}7$ ore $\rightarrow d_C = 3$ ORE COINCIDENZA PROBABILISTICA

3) ANALISI STATISTICA DEL CAMPIONE DEI MASSIMI ANNUI GENERATI

Riprendiamo l'inferenza statistica delle portate vista però in forma semplificata.

SEMPLIFICAZIONI

1 - UTILIZZARE LA DISTRIBUZIONE DI GUMBEL

Non devo usare differenti distribuzioni perché abbiamo molti dati (passate esperienze) che ci dicono che la distribuzione di Gumbel approssima bene le precipitazioni.

- **PRECIPITAZIONE DI PROGETTO**

x_T = precipitazione di durata prefissata (3h) con tempo di ritorno T in anni.

$$P = 1 - \frac{1}{T} = e^{-e^{-\frac{x-\theta_1}{\theta_2}}}$$

$$x_T = \hat{\theta}_1 - \hat{\theta}_2 \ln\left(-\ln\left(1 - \frac{1}{T}\right)\right) =$$

$$\boxed{x_T = \bar{x} - \frac{\sqrt{6}}{\pi} s_x \left[\gamma_E + \ln\left(-\ln\left(1 - \frac{1}{T}\right)\right)\right]}$$

Nel caso fortunato posso stimare x_T per ogni valore di T ($\hat{\theta}_1, \hat{\theta}_2$ dipendono dal campione).

2-B) CASO SFORTUNATO (STANDARD)

la durata critica non corrisponde a nessuna delle 5 durate fornite dal SIN.

METODO DELLA PRECIPITAZIONE INDICE

Per ognuna delle 5 durate di cui abbiamo i dati riscriviamo l'equazione:

$$x_T = \bar{x} - \frac{\sqrt{6}}{\pi} s_x \left[\gamma_E + \ln\left(-\ln\left(1 - \frac{1}{T}\right)\right)\right]$$

e successivamente interpoliamo per ottenere la precipitazione di progetto.

VARIABILE $i_T(d)$

$i_T(d)$ intensità di precipitazione corrispondente ad un certo tempo di ritorno e ad una determinata durata. $i_T = f(T, d)$.

Come si stima $\overline{CV}$?

$$\boxed{\overline{CV} = \frac{1}{5} \sum_{i=1}^{m} CV_i}$$

Inseriamo il valore medio dei 5 CV nell'equazione.

- PRECIPITAZIONE INDICE

Cerchiamo di comprendere come varia l'intensità di precipitazione media con la durata della precipitazione.

FORMULA MONOMIA $\boxed{\overline{i}(d) = a\, d^{n-1}}$

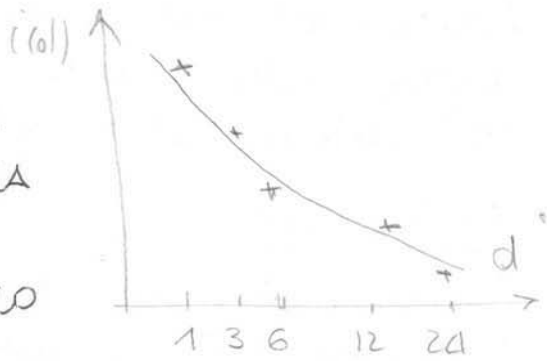
$a \simeq 15 \div 70$ mm/h^n con $d = [h]$

$n \simeq 0{,}2 \div 0{,}7$ valori tipici

n deve essere nel range di $[0,1]$ perché, matematicamente devo avere la funzione decrescente.

Perché $n \in [0,1]$?

$i(d)$ = intensità di precipitazione media, funzione della sola durata.



Per $i(d)$ dobbiamo aspettarci solo valori decrescenti col crescere della durata. VEDI PAG 39-A.

Come faccio a stimare (a, n)?

Ricorro a quei pochi dati disponibili: per quei 5 punti ho il dato campionario.

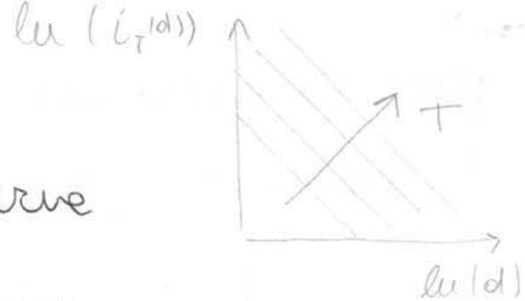
Bisogna interpolare i 5 punti con l'equazione:

$$\overline{i}(d) = a\, d^{n-1}$$

Tramite una trasformata logaritmica posso ottenere una stima lineare.

la relazione tra (i, d, T) è espressa tramite un FASCIO DI CURVE DECRESCENTI, dove la variabile tra le curve è il valore di K_T: se il tempo di ritorno cresce, K_T aumenta.
Su tempi di ritorno più grandi, avrei eventi di maggiore entità quindi, intensità di precipitazione maggiore.

$$\ln[i_T(d)] = \ln(a) + \ln(K_T) + (n-1)\ln(d)$$



Il valore di $\ln(K_T)$ varia con le curve

ALTEZZA DI PRECIPITAZIONE VS INTENSITÀ

Alcuni libri fanno riferimento alle altezze: la relazione tra queste 2 grandezze è molto semplice e si può ragionare INDIFFERENTEMENTE su una o l'altra variabile.

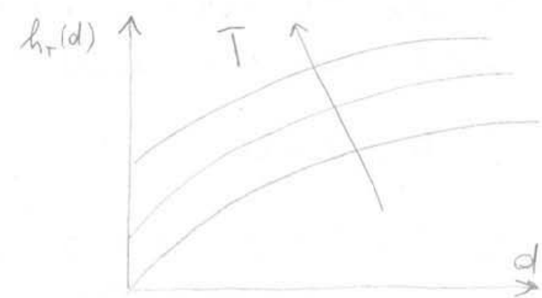
$$\boxed{h(d) = d \cdot i(d)}$$

ESEMPIO

$d = 30$ ore $\quad i(d) = 20$ mm/h $\quad h(d) = 600$ mm

$$h_T(d) = a K_T\, d \cdot d^{n-1}$$

$$\boxed{h_T(d) = a K_T d^n}$$



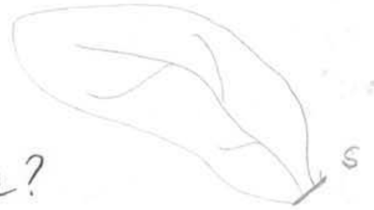
Relazione crescente con la durata (concavità verso il basso).
I dati del SIN erano riportati con altezze di precipitazione e bisognava:
- passare alle intensità $i = \frac{h}{d}$
- oppure fare il ragionamento con h e convertire dopo.

Il valore $\frac{1}{3,6}$ serve per ragionare con i m^3/s

La dipendenza dal tempo è solo funzione di $i_T(d)$ e per piccoli bacini posso ribaltare l'ipotesi di tempo di ritorno da $i_T \to Q_T$.

· OSSERVAZIONE

$$\boxed{d_c = t_{cB}}$$

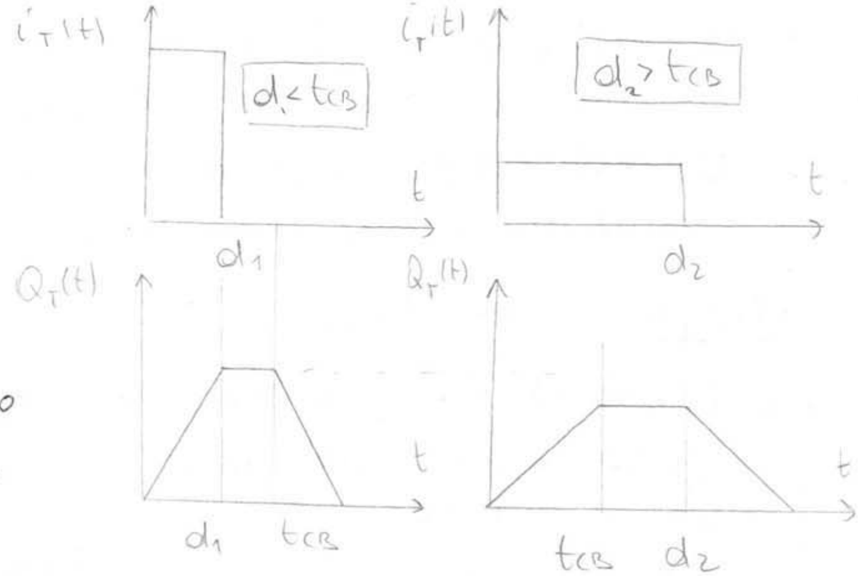


Perché avevamo fatto questa assunzione?

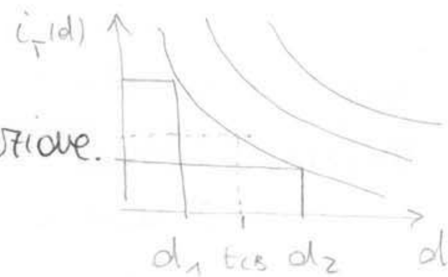
Considero 2 casi con lo stesso tempo di ritorno e:

$d_1 < t_{cB} < d_2$

Prelimiamo un certo tempo di ritorno.

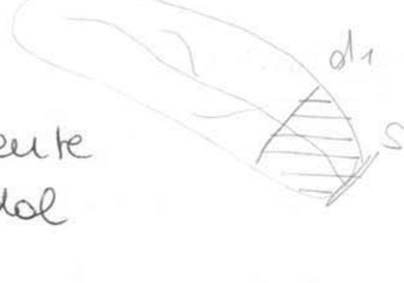


IPOTESI le intensità di precipitazione sono costanti dentro la durata dell'evento di precipitazione.



I rettangoli hanno grandezze diverse.

- CASO A: $d_1 < t_{cB}$



le portate crescono perché crescono le aree contribuenti: l'area contribuente rimane costante finché non esce dal bacino idrografico ($t = t_{cB}$).

2) BACINI MOLTO PICCOLI

Quando abbiamo durate critiche di qualche MINUTO, $d_c \to 0$, $i_T(d) \to \infty$ la formula va in crisi

$$i_T(d) = a\, K_T\, d^{n-1}$$

Quando ho durate critiche molto piccole devo usare formule diverse da quella monomia per interpolare i dati.

$$\boxed{i_T(d) = a'(1+bd)^{n'-1}}$$

b = COEFFICENTE

(a', n') hanno lo stesso significato di (a, n)

Questa formula alternativa evita che $i_T(d) \to \infty$ MA NON si può applicare sempre.

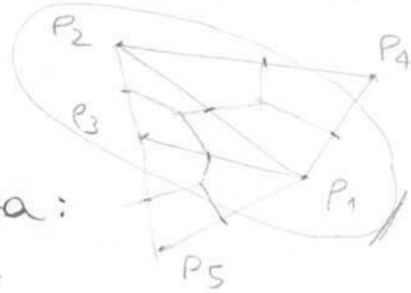
PROBLEMI

- Ho 3 parametri.
- Se usassi la trasformata logaritmica non OTTERREI una retta completa.
- E' difficile trovare i dati con tempo inferiore all'ora perché sono dati POCO SICURI.

Se le durate $t > 20 \div 30$ minuti, usiamo la formula monomia, altrimenti per $t < 20$ min. usiamo l'altra equazione.

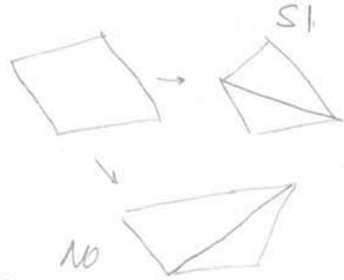
Se usassi la formula monomia per durate MOLTO BREVI incorrerei nel sovrastimare l'opera.

A) METODO DEI TOPOIETI

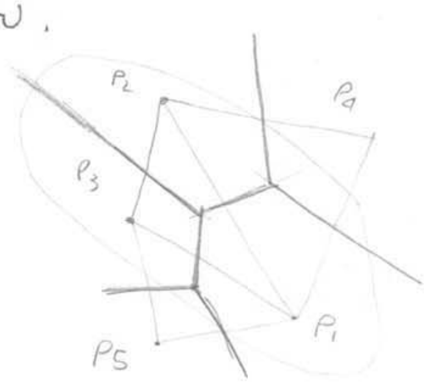


Metodo che prevede di attribuire ad ognuno dei pluviometri considerati un'area di competenza: l'ipotesi è che dentro queste aree devono ricadere punti PIÙ VICINI SOLO A QUEL pluviometro, rispetto agli altri.

- PASSAGGI

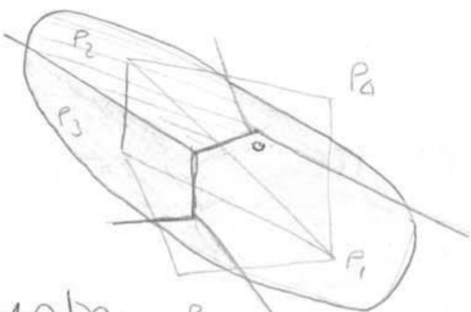


1) Unisco i pluviometri con rette, cercando di non formare triangoli con angoli TROPPO ACUTI ($\alpha \to 0$)
2) Definisco i punti medi di ogni lato e traccio le ORTOGONALI ai segmenti.
3) la divisione permette di definire le aree di competenza.



PUNTO O : punto $\in$ Area di P_1

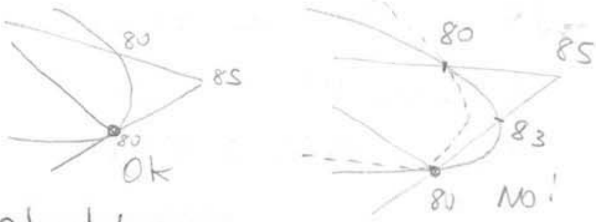
NOTA Anche i pluviometri FUORI dal bacino idrografico hanno un'area di competenza NON TRASCURABILE.



Il metodo dei Topoieti prevede di assegnare la precipitazione misurata nel pluviometro J a tutti i punti che ricadono in A_J (area di competenza del J-esimo pluviometro).

- PASSAGGI

1) Unisco i pluviometri e segno il valore P_J
2) Definisco il PASSO delle isoiete (es. 5 mm) e determino i punti sulle rette.
3) Traccio le linee con ugual valore di pioggia.
Per i punti (70, 90) ANCORA ho 1 solo valore è difficile capire cosa succede.



ATTENZIONE: non intersecare i punti di passaggio delle altre linee (questi punti potrebbero avere un valore diverso (es. 83) da quello della nostra ISOIETA (es. 80)).

Posso ricavare l'altezza di precipitazione di un qualunque punto interno al bacino

$$\boxed{P_B = \frac{1}{A_B} \iint P(\underbrace{x,y}_{\text{COORDINATE}})\, dx, dy}$$

P = PIOGGIA

- VANTAGGI

1) Costruisco un solido di forma più ragionevole, senza discontinuità.

- SVANTAGGI

1) Esiste un forte ELEMENTO DI SOGGETTIVITÀ nel definire le SPLINE, bisogna definirle bene;
2) A differenza dei Topoieti, le ISOIETE vanno tracciate ogni volta che si considera un diverso evento (pioggia minima, annua...)
→ complicazione importante.

LA PRECIPITAZIONE MEDIA

$$\boxed{P_B = \frac{1}{A_B} \iint P(x,y)\, dx\, dy}$$

- VANTAGGI

1) Metodo più sicuro da errore perchè non ha i problemi di IRREALISTICITÀ dei TOPOIETI.
2) Metodo che non ha la SOGGETTIVITÀ delle ISOIETE però non è un metodo perfetto.

- SVANTAGGI

1) L'ipotesi che la precipitazione dipenda in maniera inversamente proporzionale di quella misurata negli altri pluviometri è troppo FORZATA.

Bisogna fare le verifiche con tutti i 3 metodi.

In generale, il metodo delle distanze inverse è un buon compromesso tra SEMPLICITÀ DI APPLICAZIONE e SENSATEZZA DEI VALORI.

PROCESSI CHE PORTANO ALLA FORMAZIONE DELLE PRECIPITAZIONI

L'aria atmosferica è una miscela di diversi gas che contiene il VAPOR D'ACQUA (acqua in forma gassosa).

1) CONDENSAZIONE

La condensazione è il passaggio dallo stato gassoso allo stato liquido: per avere questo processo, la tensione di vapore dell'acqua deve essere uguale alla tensione di vapore del vapor d'acqua.

Come quantificare il vapor d'acqua?

CATENA DI EVENTI CHE PORTANO A CONDENSAZIONE

- Diminuzione di T_A ↓
- Diminuzione di e_s, e rimane costante ↓
- Incremento di RH (l'aria diventa più umida) fino al valore di RH = 1.
  Esiste una temperatura tale per cui RH = 1. ↓
- Se continuiamo ad abbassare la temperatura, una parte delle particelle del vapor d'acqua condensa e passa allo stato liquido.
- Quando RH = 1 INIZIA LA CONDENSAZIONE.

Le particelle vanno a formare l'AEROSOL ATMOSFERICO, particelle d'acqua in forma liquida di dimensioni microscopiche (come la nebbia): gocce d'acqua che galleggiano nell'atmosfera.

Perché la nebbia si forma di sera?
Di sera la T_A diminuisce, RH → 1 e le particelle cominciano a condensare.

2) COALESCENZA

La nebbia e le nuvole sono formate da gocce d'acqua sospese nell'aria: avere condensa è condizione necessaria ma non sufficiente per avere la pioggia.

Le particelle in forma liquida devono aumentare le loro dimensioni per vincere le forze di galleggiamento e cadere a terra.