Modello dell'esperimento

Sommario della lezione 3

- Modello degli esperimenti
  - Modello del processo
  - Modello per l'errore casuale
- Teoria della probabilità
- La Variabile aleatoria come modello di un esperimento
- Tipi di Variabile aleatoria
- Trasformazioni di variabili aleatorie

Modello dell'esperimento

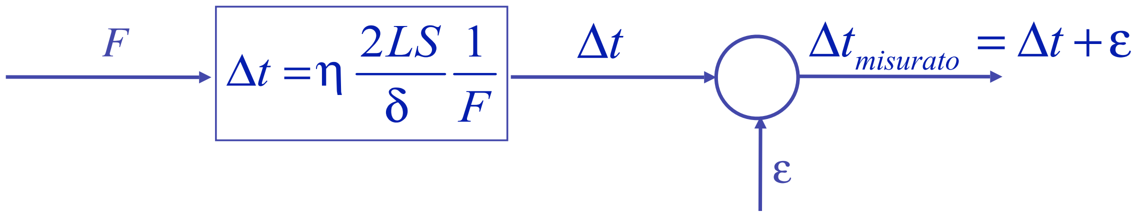
- Come abbiamo già visto il modello di un esperimento può essere schematizzato con il seguente schema a blocchi

- Una descrizione di questo tipo presume che:
  - Le condizioni sperimentali siano fissabili con precisione assoluta
  - L'errore casuale si sovrapponga alla risposta deterministica del sistema.

- Per semplicità assumiamo di essere capaci di descrivere esattamente il comportamento del processo (ipotesi molto forte).

Modello dell’esperimento

- Nell’esempio che abbiamo già considerato del viscosimetro



- E’ evidente che l’errore casuale gioca un ruolo importante.

- Se non ci fosse errore casuale ad ogni replica dell’esperimento otterremmo esattamente lo stesso risultato.
  - Abbiamo già scoperto invece che in qualunque esperimento reale sono presenti errori casuali

Errori casuali

- Cosa sappiamo dell'errore?

- Sebbene le condizioni sperimentali siano note perfettamente, e sebbene stiamo considerando il modello del processo vero, non siamo capaci di prevedere il risultato dell'esperimento.

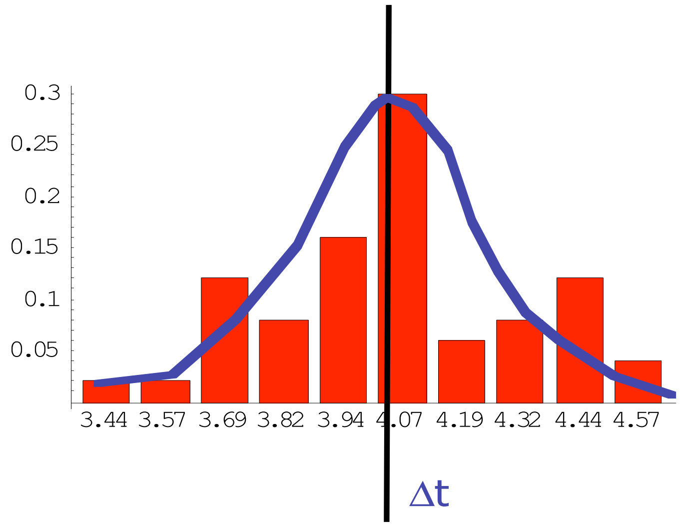
- A valle della sperimentazione disponiamo dei diagrammi di frequenza. Se il modello del processo che stiamo considerando è "vero" possiamo usare i dati sperimentali per imparare qualcosa sull'errore.

$$\Delta t_{misurato} = \Delta t + \varepsilon$$
$$\Delta t_1 = \Delta t + \varepsilon_1$$
$$\Delta t_2 = \Delta t + \varepsilon_2$$
$$\dots$$
$$\Delta t_N = \Delta t + \varepsilon_N$$

Errori casuali

- Ad ogni replica il risultato dell'esperimento si compone di una parte che non cambia Δt (ma che non conosciamo) ed una parte che "peschiamo" a caso.

- Questa parte "variabile" è la causa del tipo di frequenza che osserviamo.



Errori casuali

- Se così stanno le cose l'errore casuale ha delle proprietà. E come se pescassimo a caso da una distribuzione di probabilità che ha una forma specifica.

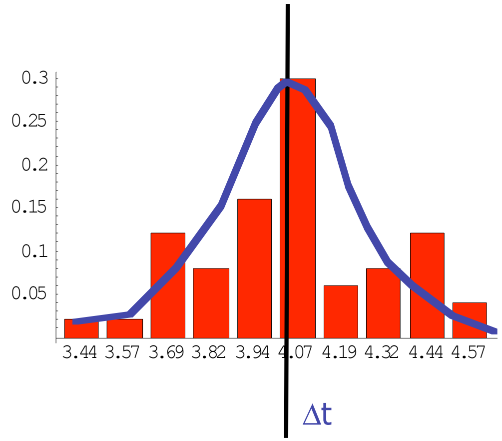
- Cosa sappiamo dell'errore?

- Ipotesi:
  - Il modello del processo che stiamo considerando è "vero"
  - Non commettiamo errori sistematici nella sperimentazione
  - Facciamo molte repliche

- Usiamo i dati sperimentali.

Errori casuali

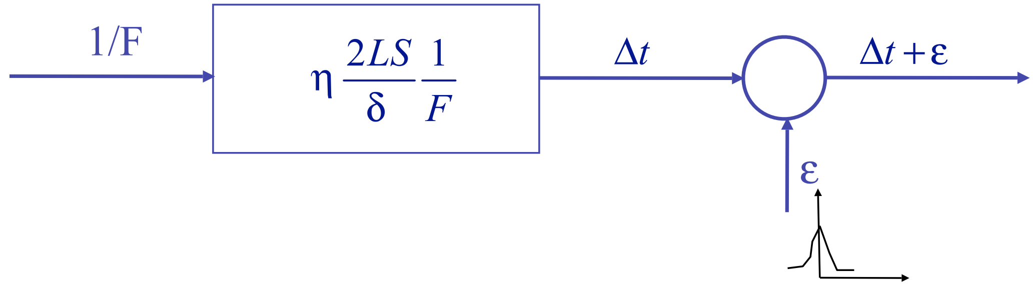
- La forma dell'istogramma della frequenza ci ricorda una curva di densità Normale…



- Potremmo forse cominciare a convincerci che l'errore sperimentale sia una grandezza caratterizzata da una distribuzione a campana con media zero (perchè?).

- Quindi **l'errore di misura può essere modellato**.

Modello dell'esperimento



$$\underbrace{Y}_{?} = \underbrace{f\left(X\right)}_{\textit{Deterministico}} + \underbrace{\varepsilon}_{\textit{Aleatorio}}$$

Errori casuali

- **Il modello dell'esperimento è la distribuzione di probabilità dello spazio campionario.**

- Questo modello conterrà parametri
  - Media, deviazione standard…

- L'identificazione parametrica porterà alla conoscenza di queste informazioni.

- E' evidente quanto sia cruciale la conoscenza della teoria della probabilità.

Probabilità

- La teoria della probabilità nasce con la necessità di prevedere i risultati di giochi d'azzardo.
  - 1657: Christiaan Huygens pubblica De ludo aleae, il primo libro sulla probabilità
  - 1708: Pierre de Montmort scrive Essai d'analyse sur le jeux de hasard.
  - 1713: viene pubblicato postumo Ars conjectandi di Jakob Bernoulli dove tra l'altro viene formulato il primo teorema limite, ovvero la legge dei grandi numeri.

- Il più semplice gioco è il lancio di una moneta
  - Intuitivamente sappiamo che la probabilità di avere testa o croce sono uguali
  - Possiamo assegnare una probabilità 0.5 a ciascun evento

- In genere la probabilità di tutti i possibili eventi è posta pari ad 1

- Il lancio di una moneta è un esperimento il cui esito è incerto.

- Intuitivamente noi presumiamo che tutti i lanci siano estratti da una popolazione sottostante in cui la probabilità di avere testa è 0.5 e croce 0.5.
  - Però se lanciamo la moneta 100 volte possiamo avere 54 teste e 46 croci
- **Non possiamo verificare esattamente la nostra stima intuitiva, ma con campioni molto grandi possiamo avvicinarci ad essa**.

Teoria della Probabilità (a chiacchiere!)

- La probabilità è un numero compreso tra zero ed uno che esprime quanto verosimilmente un evento può accadere
- La probabilità di composizioni di eventi non collegati tra loro è la somma delle probabilità dei singoli eventi.
- Si possono calcolare le probabilità di eventi che si verificano se altri eventi si sono verificati.

Teoria della Probabilità (rigorosa!)

- Si chiama funzione di probabilità una funzione di insieme a valori reali definita su R avente le seguenti proprietà:

ASSIOMI (Kolmogoroff, 1933)

1) $0 \le P(E) \le 1$ Gli eventi spaziano dal non accadere mai all'accadere sempre

2) $P(\Omega) = 1$ Qualcosa deve accadere

3a) $Se\ E_1\ ed\ E_2\ sono\ disgiunti$
$allora\ P(E_1 \cup E_2) = P(E_1) + P(E_2)$

3b) $Se\ E_1\ E_2, \dots\ sono\ a\ due\ a\ due\ disgiunti,\ allora$
$P(E_1 \cup E_2 \dots \cup E_n) = P(E_1) + P(E_2) + \dots + P(E_n)$

Probabilità condizionata

- Consideriamo l'evento L: una persona è nata in un mese di 31 giorni
- Consideriamo l'evento R: una persona è nata in un mese il cui nome contiene la lettera r
- Le probabilità di tali eventi sono facili da calcolare:
  - L = {Gen, Mar, Mag, Lug, Ago, Ott, Dic} P(L)=7/12
  - R = {Feb, Mar, Apr, Set, Ott, Nov, Dic} P(R)=7/12
- Supponiamo di sapere che la persona in questione è nata in un mese di 31 giorni, quale è la probabilità che sia nata in un mese il cui nome contiene la lettera r?
  - L'informazione che abbiamo esclude 5 esiti del nostro spazio campionario: la persona non può essere nata a Feb, Apr, Giu, Sett, o Nov. Sono possibili solo 7 esiti di cui solo 3 cadono in R ∩ L = {Mar, Ott, Dic} quindi la probabilità è 3/7
- Nel gergo statistico questo tipo di probabilità viene espressa con la notazione:

$$P(R|L) = \frac{3}{7}$$

Questo non equivale alla P(R∩L), che è pari a 1/4.
Notate anche che P(R|L) è la frazione P(R∩ L)/P(L).

Probabilità Condizionata

- Supponete di essere interessati solo ad una parte dello spazio campionario, la parte in cui voi sapete che un evento A si è verificato e volete conoscere con quale probabilità altri eventi in esso contenuti (in A) si verifichino.
- **Probabilità condizionata** che si verifichi B se A si è verificato

Probabilità condizionata

$$P(B|A)=\frac{P(A\cap B)}{P(A)} \qquad (1)$$

$$P(A|B)=\frac{P(A\cap B)}{P(B)} \qquad (2)$$

  - La (1) vale purché $P(A)\neq 0$ e la (2) purché $P(B)\neq 0$
  - Se $B\subset A$ allora $P(A|B)=1$
  - Se $A\subset B$ allora $P(A|B)=P(A)/P(B)\geq P(A)$
  - Le probabilità condizionate sono probabilità perché soddisfano gli assiomi

Probabilità Condizionata

- Tempi di residenza in un reattore CSTR
  - Consideriamo un CSTR che opera in fase liquida a perfetta miscelazione
  - La distribuzione dei tempi di residenza ci dice quanto a lungo particelle di liquido restano nel reattore prima di uscirne.
  - Sia Rt l'evento "la particella ha tempo di residenza maggiore di t secondi"
  - Dal corso di Progettazione sapete che la distribuzione dei tempi di permanenza è e^{-t} .
  - Quindi:

$$P(R_3) = e^{-3} = 0.04978...$$

$$P(R_4) = e^{-4} = 0.01831...$$

  - Possiamo usare la definizione di probabilità condizionata per trovare la probabilità che una particella che è rimasta nel reattore 3 secondi ci rimarrà per più di 4 secondi.

$$P(R_4|R_3) = \frac{P(R_4 \cap R_3)}{P(R_3)} = \frac{P(R_4)}{P(R_3)} = \frac{e^{-4}}{e^{-3}} = 0.36787...$$

Teoria della Probabilità (rigorosa!)

INDIPENDENZA STOCASTICA

- La nozione di indipendenza stocastica di eventi è fondamentale nella teoria della probabilità e nella pratica della sperimentazione.

- Se due eventi in un esperimento aleatorio sono tali da verificare che

  $P(A \cap B)=P(A)\ P(B)$ allora essi si dicono indipendenti.

- Dalla definizione di probabilità condizionata $P(A \cap B)=P(B)P(A|B)$.

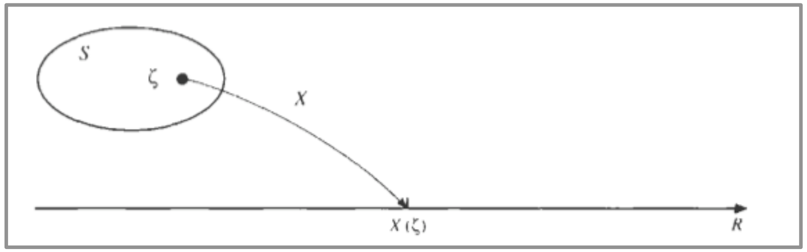
- Nel caso in cui $P(A \cap B)=P(A)P(B)$ si ottiene: $P(A|B)=P(A)$.
  - **Qualunque cosa accada a B essa non dà informazioni su A. Quindi A e B sono indipendenti**

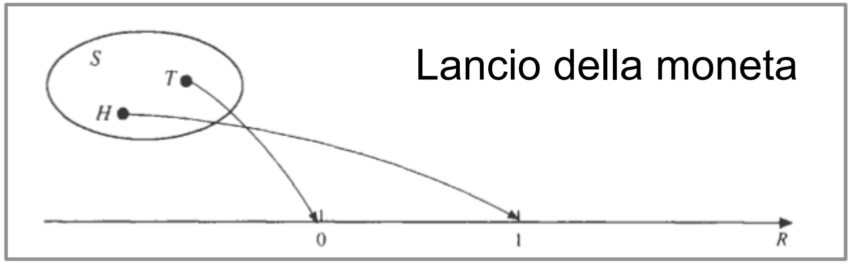
LA VARIABILE ALEATORIA

- Consideriamo il lancio di una moneta ed assegnamo il valore 1 alla testa e 0 alla croce. Segnaliamo il possibile risultato con la lettera Y. Prima di lanciare la moneta ad Y non può essere assegnato un valore definito
  - Potrà assumere valore 1 con probabilità ½ (se esce testa) oppure 0 con probabilità ½ se esce croce.
- Può essere ragionevole chiamare Y **Variabile Aleatoria** (VA), essa può essere pensata come il risultato numerico di un esperimento quando questo non è prevedibile con certezza (ossia non è deterministico).
  - In generale, una VA è una variabile il cui valore non è fissato, ma può assumerne uno qualunque preso da un intervallo di valori con probabilità prefissata.
  - Per caratterizzare una VA completamente si dovrebbe conoscere non solo quali valori possa assumere ma anche quanto frequentemente li assuma, ovvero quale probabilità abbia di assumerli.
- **Possiamo modellare il risultato di un esperimento come una VA.**

La variabile aleatoria

- Consideriamo un esperimento aleatorio con spazio campionario Ω. Una variabile aleatoria X(ζ) è una funzione reale che assegna un numero reale (detto il valore di X(ζ)) ad ogni punto ζ di Ω.
  - Chiaramente una variabile aleatoria è una funzione.
- Lo spazio campionario Ω è il dominio della VA, e la collezione di tutti i numeri (i valori di X(ζ)) è il codominio di X.
- Quindi il codominio di X è un certo sottoinsieme dei numeri reali





La variabile aleatoria

- Se X è una VA ed x un numero reale, possiamo definire l'evento (X=x) come:

$$(X = x) = \{\zeta : X(\zeta) = x\}$$

- Altri esempi

$$(X \le x) = \{\zeta : X(\zeta) \le x\}$$
$$(X > x) = \{\zeta : X(\zeta) > x\}$$
$$(x_1 < X \le x_2) = \{\zeta : x_1 < X(\zeta) \le x_2\}$$

- Probabilità associate

$$P(X = x) = P\{\zeta : X(\zeta) = x\}$$
$$P(X \le x) = P\{\zeta : X(\zeta) \le x\}$$
$$P(X > x) = P\{\zeta : X(\zeta) > x\}$$
$$P(x_1 < X \le x_2) = P\{\zeta : x_1 < X(\zeta) \le x_2\}$$

LA VARIABILE ALEATORIA

- In genere le VA sono rappresentate da lettere maiuscole.

- Il risultato (esito) di un esperimento è rappresentato da minuscole

- Noi useremo questa nomenclatura: "*Il lancio di questo dado è una variabile aleatoria Y; il risultato di questo lancio è y=5*"
  - Si noti che Y è una funzione (variabile aleatoria) mentre i valori assunti da tale funzione, $y = Y(\omega)$, valori calcolabili quando l'esito dell'esperimento sia noto, ovvero sono numeri reali.
  - Fidatevi che questa distinzione sulla nomenclatura genererà chiarezza!

LA VARIABILE ALEATORIA

- In genere ci si porrà domande del tipo:
  - Qual è la probabilità che Y sia minore di a?
  - Qual è la probabilità che Y sia compresa tra a e b?

- Le probabilità si assegnano ad eventi quindi condizioni su Y devono corrispondere ad eventi. Da ora in poi, eventi saranno indicati come segue:

- $\{Y<y\}$: Sottoinsieme di Ω che contiene risultati y_i tali che $Y(y_i)<y$.

- Esempi:
  - $\{y_1 \leq Y \leq y_2\}$,
  - $\{Y=y\}$,
  - $\{Y \in \mathrm{R}\}$

LA VARIABILE ALEATORIA

- Gli elementi di Ω che sono contenuti nell'evento $\{Y \leq y\}$ cambiano al variare di y.
  - La probabilità $P\{Y \leq y\}$ dipende quindi da y.
- Questa funzione la possiamo definire brevemente come: $F(y)=P\{Y \leq y\}$ che chiameremo funzione di distribuzione della probabilità o funzione cumulativa della distribuzione di probabilità (**CDF**), o funzione di ripartizione
- Si può dimostrare che per ogni variabile aleatoria Y la funzione (reale di una variabile reale) distribuzione di probabilità $F_Y(y) = P\{Y \leq y\}$ è sufficiente per definire la probabilità di qualunque evento.
- Tale funzione è utile per il calcolo della probabilità di un evento qualunque ma in molte circostanze è più comodo l'uso della sua derivata $f_Y(y)$, la **funzione densità di probabilità** (PDF), che ha spesso espressione analitica, anche se non sempre semplice.

Funzioni distribuzione e densità di probabilità

- La relazione tra CDF e PDF è esprimibile con le relazioni

$$F(y) = \int_{-\infty}^{y} f(u)\,du \qquad f(y) = \frac{dF(y)}{dy}$$

- La PDF è evidentemente una funzione non negativa ed è tale che

$$\int_{-\infty}^{\infty} f(u)\,du = 1$$

- Nei casi di maggiore interesse in questa sede la funzione $F(y)$ è una funzione continua (variabili aleatorie continue) e , quindi, $P\{\omega\}$ = 0 per ogni $\omega \in \Omega$, con ω l'evento elementare

CDF

- Sia: $F\left(y^{+}\right)=\lim_{0<\varepsilon\to 0} F\left(y+\varepsilon\right)$

  $F\left(y^{-}\right)=\lim_{0<\varepsilon\to 0} F\left(y-\varepsilon\right)$

PROPRIETA' DI F

1. $F(+\infty)=1$, $F(-\infty)=0$, infatti:
   - $F(+\infty)=P\{Y\leq +\infty\}=P\{\Omega\}=1$
   - $F(-\infty)=P\{Y\leq -\infty\}=0$

- **F è una funzione non decrescente**

- Se $F(y_0)=0$ allora $F(y)=0$ se $y<y_0$

CDF

PROPRIETA' DI F

4. $P\{Y>y\} = 1- F(y)$, infatti:

   $\{Y>y\}$ e $\{Y\leq y\}$ sono disgiunti per cui $\{Y>y\}\cup\{Y\leq y\}=\Omega$
   da cui deriva la proprietà.

5. F è una funzione continua da destra: $F(y^+)=F(y)$

6. $P\{y_1<Y\leq y_2\}= F(y_2)-F(y_1)$

   $\{Y\leq y_1\}$ e $\{y_1<Y\leq y_2\}$ sono disgiunti.

   D'altra parte, $\{Y\leq y_2\}= \{Y\leq y_1\} \cup\{y_1<Y\leq y_2\}$ da cui si deduce la proprietà.

- **Dunque nota la $F(y)$ siamo in grado di ricavare la probabilità di qualunque evento**

PROPRIETA' DI F

7. $P\{Y = y\} = F(y)\text{-}F(y^{-})$

- La variabile aleatoria è continua se F è continua. Si deduce quindi che in questo caso l'evento elementare ha probabilità nulla, infatti se F è continua $F(y)=F(y^{-})$ e $P\{Y = y\}$ = 0.

- La variabile aleatoria è discreta se F è a gradini. In questo caso l'evento elementare ha probabilità finita: $F(y_i)\text{-}F(y_i^{-})=P\{Y=y_i\}=p_i$

FUNZIONE DENSITA' DI PROBABILITA'

- Abbiamo definito funzione densità di probabilità della variabile aleatoria la funzione:

$$f(y) = \frac{d}{dy} F(y) \quad se\, la\, VA\, è\, continua$$

$$f(y) = \sum_i p_i \delta(y - y_i) \quad se\, la\, VA\, è\, discreta$$

PROPRIETA'

1) $f(y) \geq 0$

2) $F(y) = \int_{-\infty}^{y} f(\xi)\, d\xi$

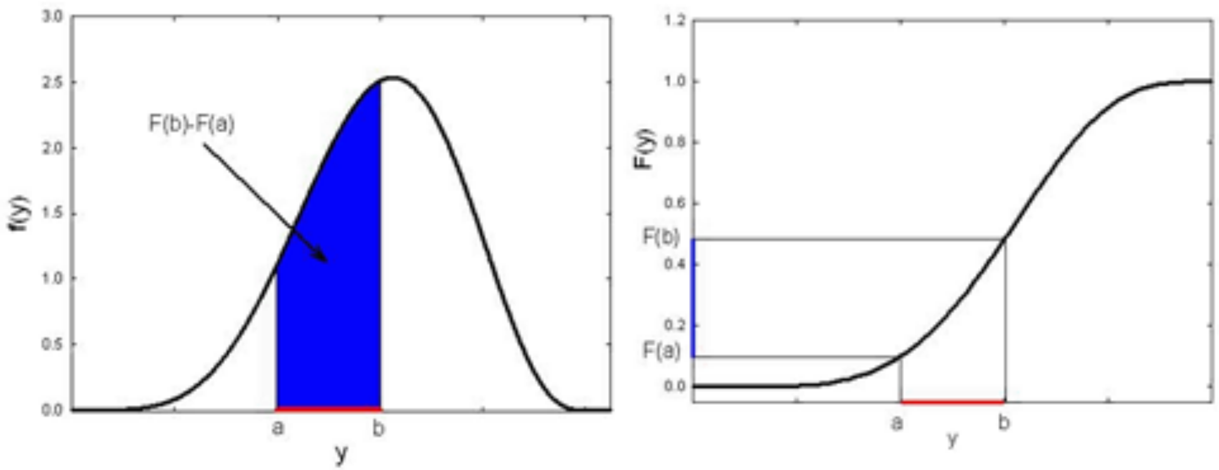
3) $\int_{-\infty}^{+\infty} f(\xi)\, d\xi = 1$

4) $F(y_2) - F(y_1) = \int_{y_1}^{y_2} f(\xi)\, d\xi$

CDF e PDF

- L'uso della CDF per il calcolo della probabilità è più semplice dell'uso della pdf che richiede il calcolo di un integrale!

- Il calcolo della probabilità $P\{a \leq Y \leq b\}$ con la PDF e la CDF è mostrato nelle figure seguenti



CDF e PDF

- La funzione $f(y)$ prende anche il nome di **funzione verosimiglianza** di Y.

- La probabilità che la variabile aleatoria Y capiti in un piccolo intervallo I di centro y ed ampiezza Δ (costante e "piccola") dipende dal valore y ed è approssimativamente dato da

$$P(I) = f(y)\,\Delta$$

STATISTICA DESCRITTIVA

- Spesso, è sufficiente per descrivere una distribuzione (almeno in forma approssimata) conoscere alcune grandezze che la caratterizzano come ad esempio media e varianza.

- **Il valore medio** di una distribuzione è dato da:

$$\mu = \sum_j y_j \, f\left(y_j\right) \qquad \textit{Caso discreto}$$

$$\mu = \int_{-\infty}^{+\infty} y \, f\left(y\right) dy \qquad \textit{Caso continuo}$$

- Se una distribuzione è simmetrica rispetto a $y=c$ allora:
  - $f(c+y)=f(c-y)$
  - $\mu=c$.

STATISTICA DESCRITTIVA

- **La varianza** di una distribuzione è:

$$\sigma^2 = \sum_j \left(y_j - \mu\right)^2 f\left(y_j\right) \qquad \textit{Caso discreto}$$

$$\sigma^2 = \int_{-\infty}^{+\infty} \left(y - \mu\right)^2 f\left(y\right) dy \qquad \textit{Caso continuo}$$

  – In genere $\sigma^2>0$

- **Valore atteso**:

$$\mathrm{E}\left(g\left(Y\right)\right) = \sum_j g\left(y_j\right) f\left(y_j\right) \qquad \textit{Caso discreto}$$

$$\mathrm{E}\left(g\left(Y\right)\right) = \int_{-\infty}^{+\infty} g\left(y\right) f\left(y\right) dy \qquad \textit{Caso continuo}$$

  – $g(Y)$ deve essere definita per ogni Y per cui $f(y)$ non è nulla. (**N.B. abbiamo introdotto una funzione di una variabile aleatoria.**

STATISTICA DESCRITTIVA

- Media e varianza sono valori attesi:

$$\mu = \mathrm{E}(Y) = \int_{-\infty}^{+\infty} y\, f(y)\, dy$$

$$\sigma^2 = E\left((Y-\mu)^2\right) = \int_{-\infty}^{+\infty} (y-\mu)^2 f(y)\, dy$$

- Possiamo generalizzare media e varianza con i **momenti**

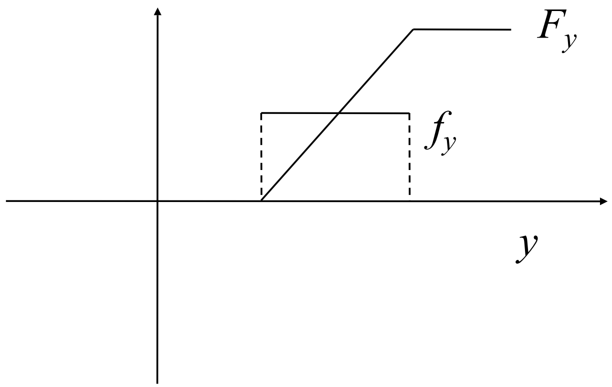
$$m_n = \int_{-\infty}^{+\infty} y^n f(y)\, dy \qquad momento\ n-simo$$

$$M_n = \int_{-\infty}^{+\infty} (y-\mu)^n f(y)\, dy \qquad momento\ centrale\ n-simo$$

UNIFORME

- Una variabile aleatoria si dice uniforme tra y_1 e y_2 se la sua densità di probabilità è costante in (y_1, y_2)

$$\boxed{f_Y = \begin{cases} \dfrac{1}{y_2 - y_1} & se \;\; y_1 \le y \le y_2 \\ 0 & altrimenti \end{cases}}$$



- **La distribuzione dipende da due parametri**

TIPI DI V.A.

NORMALE o GAUSSIANA

- Una variabile aleatoria si dice normale o Gaussiana, Y≡N(μ, σ), se la densità di probabilità è:

$$f_Y = \frac{1}{\sqrt{2\pi}\ \sigma_Y} \exp\left[-\frac{1}{2} \frac{\left(y-\mu_Y\right)^2}{\sigma_Y^2} \right]$$

- Il grafico di tale funzione è una curva a campana simmetrica rispetto a $y=\mu_Y$

- Anche in questo caso **la distribuzione dipende da due parametri**

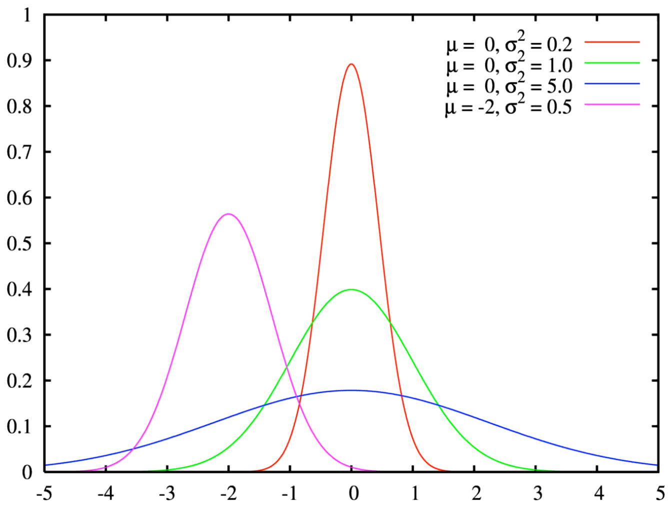
- **La maggior parte delle variabili aleatorie con cui si avrà a che fare sono Gaussiane**.

STORIA (da wikipedia)
Karl Friedrich Gauss descrisse la Normale studiando il moto dei corpi celesti. Altri la usavano per descrivere fenomeni anche molto diversi come i colpi di sfortuna nel gioco d'azzardo o la distribuzione dei tiri attorno ai bersagli. Da qui i nomi curva i Gauss e curva degli errori:
Nel 1835 Lambert-Adolphe-Jacques Quételet pubblicò uno scritto nel quale, fra le altre cose, c'erano i dati riguardanti la misura del torace di soldati scozzesi e la statura dei militari di leva francesi. Quételet mostrò come tali dati si distribuivano come una Gaussiana, ma non andò oltre.
Fu Francis Galton a intuire che la curva in questione, da lui detta anche ogiva, poteva essere applicata a fenomeni anche molto diversi, e non solo ad "errori". Questa di idea di curva per descrivere i "dati" in generale portò ad usare il termine Normale, in quanto rappresentava uno substrato normale ovvero la norma per qualsiasi distribuzione presente in natura.
Nel tentativo di confrontare curve diverse, Galton - in mancanza di strumenti adeguati - si limitò ad usare due soli parametri: la media e la varianza, dando così inizio alla statistica parametrica.

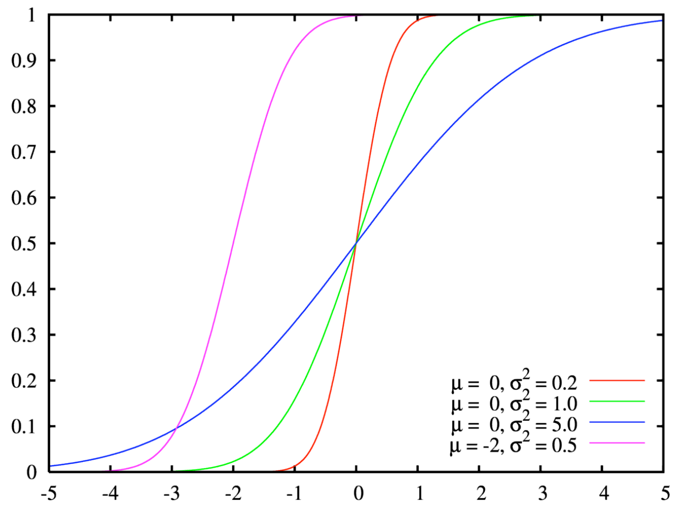
NORMALE o GAUSSIANA

- Non è possibile calcolare in forma analitica la corrispondente funzione distribuzione (CDF) (è calcolabile da tabelle)

$$F_Y = -\frac{1}{2}\left(1 + Erf\left[\frac{(y - \mu_Y)}{\sqrt{2}\sigma_Y}\right]\right)$$



PDF normale



CDF normale

Trasformazioni di una VA

Dipendenza della viscosità dalla temperatura

- La viscosità di un liquido omogeneo Newtoniano dipende fortemente dalla temperatura.
- Al crescere della temperatura la viscosità diminuisce.
- Una relazione empirica che modella la dipendenza della viscosità dalla temperatura è alla Arrhenius

$$\eta = \eta_0 \, e^{\frac{E}{RT}}$$

- Se torniamo all'esempio del viscosimetro a banda dobbiamo modificare un po' le cose per tener conto della dipendenza da T della viscosità

Trasformazioni di una VA

- L'esperimento è condotto a T costante (che immaginiamo di saper fissare con la precisione che vogliamo)

$$\Delta t = \eta_0 \, e^{\frac{E}{RT}} \frac{2LS}{\delta} \frac{1}{F}$$

  - Abbiamo due variabili indipendenti: F e T
  - Durante la campagna sperimentale misureremo per varie scelte di F e T il valore dell'intervallo di tempo (VA).

- Ora però il modello è nonlineare nel parametro E.
  - Come vedremo più avanti questo aspetto complica un po' il processo di stima dei parametri.

- Con un po' di intuizione matematica possiamo però linearizzare il modello

Trasformazioni di una VA

- Applichiamo la funzione Logaritmo ad entrambi i membri dell'equazione del viscosimetro:

$$Log(\Delta t) = Log\left(\eta_0\, e^{\frac{E}{RT}} \frac{2LS}{\delta}\frac{1}{F}\right)$$

$$Log(\Delta t) = Log(\eta_0) + \frac{E}{RT} + Log\left(\frac{2LS}{\delta}\frac{1}{F}\right)$$

$$y = a + bx_1 + x_2 \qquad\qquad con\;\; x_1 = \frac{1}{T} \;\; e \;\; x_2 = \frac{2LS}{\delta}\frac{1}{F}$$

- Siamo arrivati ad un modello lineare nei parametri applicando una trasformazione nonlineare al modello originario.
  - Abbiamo operato sul modello del processo
  - Cosa succede se operiamo sul modello dell'esperimento?

Trasformazioni di una VA

- **Modello dell'esperimento**
  - La VA è Δt

$$\Delta t_{misurato} = \eta_0 \, e^{\frac{E}{RT}} \frac{2LS}{\delta} \frac{1}{F}$$

$$Log\left(\Delta t + \varepsilon\right) = Log\left(\eta_0 \, e^{\frac{E}{RT}} \frac{2LS}{\delta} \frac{1}{F}\right)$$

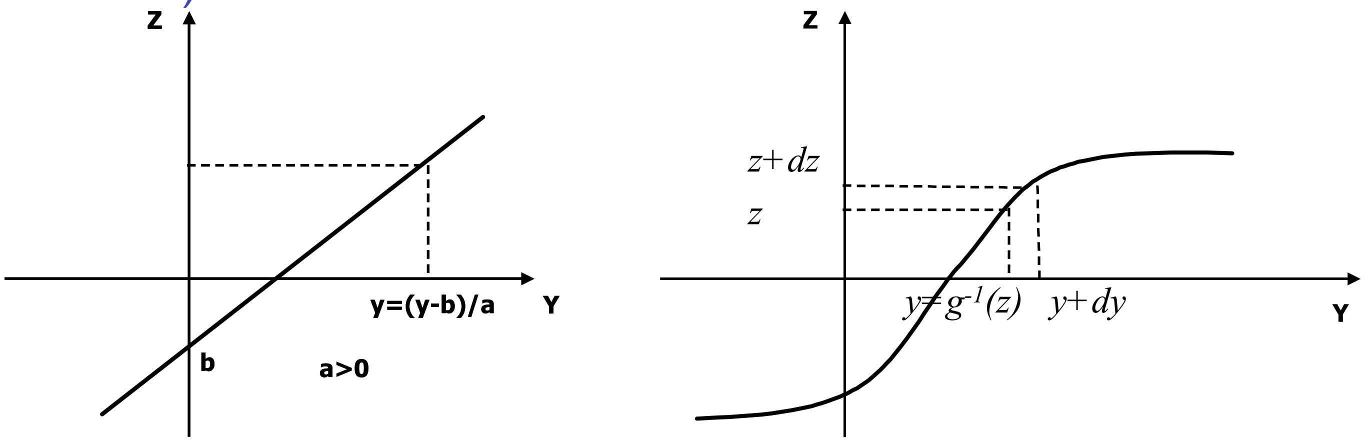
- Quando applichiamo il Logaritmo trasformiamo la VA

- L'operazione di trasformazione ha un effetto sulle proprietà della VA?

Trasformazioni di una VA

- In molti casi si presentano situazioni in cui a partire da una variabile aleatoria Y si passa una nuova variabile aleatoria Z mediante una opportuna trasformazione.

- Esempio nell'esperimento si misura una concentrazione e si calcola la conversione (con trasformazioni lineari) o il logaritmo (trasformazione non lineare) della concentrazione stessa.



Trasf. Lineare Z = aY + b　　　**Trasformazione non lineare** Z = g(Y)

Trasformazioni di una VA

- In definitiva Z è la variabile aleatoria definita sullo spazio di probabilità (Ω, Ω_E, P) dalla funzione composta g (Y)

- Si può determinare la funzione densità di probabilità della variabile aleatoria Z in funzione di quella di Y?

- La VA Z è dello stesso tipo della VA Y ? Cioè il "tipo" della pdf è lo stesso, a meno del valore dei parametri eventualmente presenti?

- Se la trasformazione è lineare il tipo di VA non cambia con la trasformazione ma cambia il valore dei parametri che la caratterizzano.

- Se la trasformazione è non lineare cambia il tipo di VA

FUNZIONI DI UNA V.A.

- Sia Y una variabile aleatoria e sia g(y) una funzione reale della variabile reale y.

$$Z = g(Y)$$

- Z è una variabile aleatoria, infatti per un dato ξ, Y(ξ) è un numero e g[Y(ξ)] è un altro numero specificato attraverso Y(ξ) e g(y).

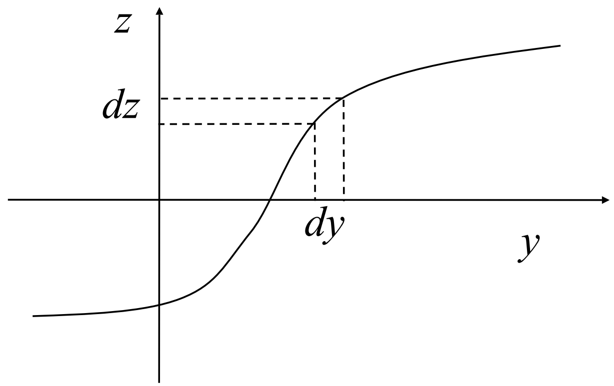
- Dato che Z è una variabile aleatoria essa è caratterizzata da una funzione di distribuzione:

$$F_Z(z) = P\{Z \le z\} = P\{g(Y) \le z\}$$

  – Il pedice Z è usato per riferire F alla VA Z.

FUNZIONI DI UNA V.A.

- Ci si può domandare come sia distribuita Z. Procediamo per gradi, ed supponiamo che la funzione g sia monotona in modo da avere una corrispondenza **biunivoca** tra Y e Z.



- Se conosco il risultato dell'esperimento espresso come y conosco anche la corrispondente z.

$$P\{y < Y \leq y + dy\} = P\{z < Z \leq z + dz\}$$

FUNZIONI DI UNA V.A.

- z e dz non sono qualunque ma corrispondono a y e dy

$$\begin{cases} z = g(y) \\ y = g^{-1}(z) \end{cases}$$

$$P\{y < Y \le y + dy\} = \int_{y}^{y+dy} f(\xi)\, d\xi = f_Y(y)\, dy$$

$$P\{z < Z \le z + dz\} = f_Z(z)\, dz$$

$$f_Y(y)\, dy = f_Z(z)\, dz \quad \Rightarrow \quad f_Z(z) = \frac{f_Y(y)}{\frac{dz}{dy}}$$

- **Obiettivo**: conoscere $f_Z(z)$

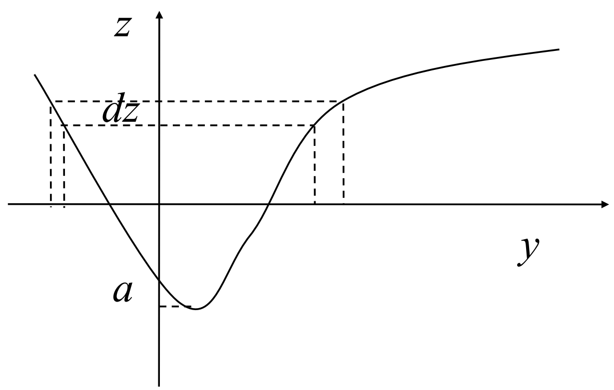
FUNZIONI DI UNA V.A.

- Procedimento
  - Fissiamo z
  - Risolviamo per y: $y=g^{-1}(z)$
  - Valutiamo $\left.\frac{dz}{dy}\right|_z = \left.\frac{dg(y)}{dy}\right|_z$
  - Quindi determiniamo

$$\boxed{f_Z(z) = \frac{f_Y(y)}{\left|\left.\frac{dz}{dy}\right|_z\right|} = \frac{f_Y\left[g^{-1}(z)\right]}{\left|g'(y)|_z\right|}}$$

FUNZIONI NON BIUNIVOCHE

- Ad un evento su z possono corrispondere più eventi su y.



- In questo esempio $f_Z(z)$ per $z<a$ è nulla

FUNZIONI NON BIUNIVOCHE

- Procedimento
  - z=g(y)
  - Si fissa z quindi si ricavano le radici y
  - P{z<Z≤z+dz}=ΣP{y<Y≤y+dy}

- Quindi anche in questo caso se si conosce f_Y possiamo ricavare f_Z anche se il procedimento è un po' più complesso

$$f_Z = \sum_{i=1}^{NR} \frac{f_Y(y_i)}{\left| \left. \frac{dz}{dy} \right|_{y_i} \right|} \qquad NR = numero\ radici$$

Trasformazioni di una variabile aleatoria: trasformazioni non biunivoche

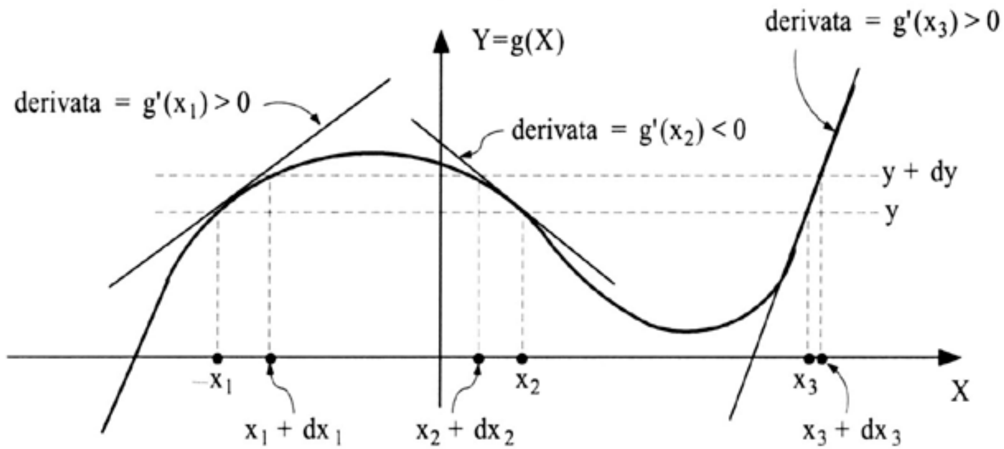


Fig. 4.14. Dimostrazione del teorema fondamentale sulle trasformazioni di variabili aleatorie. Le soluzioni dell'equazione $y = g(x)$ sono x_1, x_2, ed x_3.

FUNZIONI DI UNA V.A.

ESEMPIO

- Z=aY+b
- g'(y)=a
- L'equazione z=ay+b ha una unica soluzione: $y = \frac{z-b}{a}$
- Quindi: $f_Z = \frac{1}{|a|} f_Y\left(\frac{z-b}{a}\right)$

Trasformazione Lineare

- **Se la funzione di trasformazione è lineare non cambia il tipo di variabile aleatoria**

FUNZIONI DI UNA V.A.

ESEMPIO

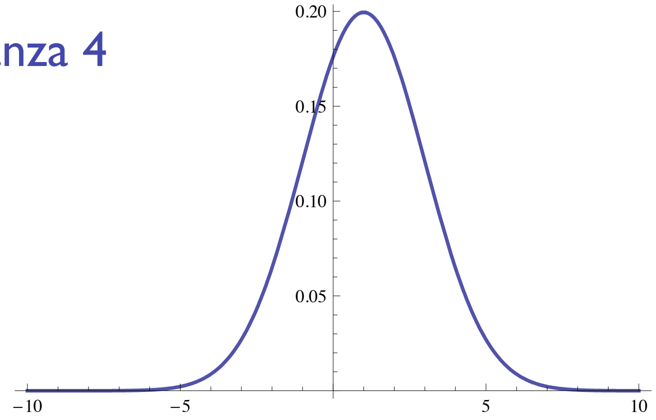
Trasformazione Non Lineare

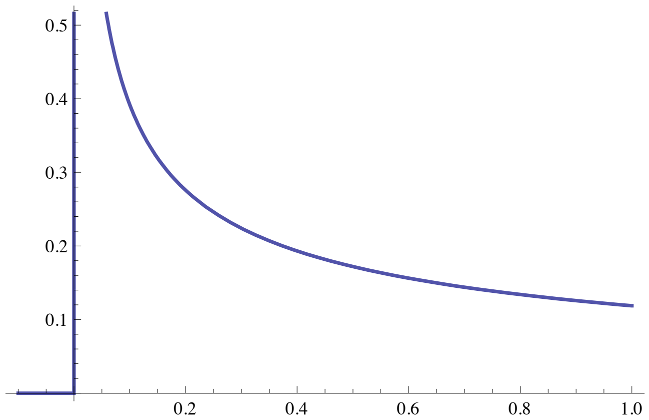
- $Z=aY^2$
- $g'(y)=2ay$
- Se $z<0$ non ci sono soluzioni reali quindi $f_Z=0$
- Se $z>0$: $y_1=\sqrt{\frac{z}{a}}, \quad y_2=-\sqrt{\frac{z}{a}}$
- Quindi: $f_Z=\frac{1}{\left|2a\sqrt{\frac{z}{a}}\right|}\left[f_Y\left(\sqrt{\frac{z}{a}}\right)+f_Y\left(-\sqrt{\frac{z}{a}}\right)\right]$
- **Se la funzione di trasformazione è non lineare, il tipo di variabile aleatoria cambia**

FUNZIONI DI UNA V.A.

ESEMPIO Y=N[1,4] Normale con media 1 e varianza 4

- $Z=2Y^2$
- $g'(y)=4y$
- Se z<0 non ci sono soluzioni reali quindi $f_Z=0$
- Se z>0: $y_1 = \sqrt{\frac{z}{2}} \quad y_2 = -\sqrt{\frac{z}{2}}$
- Quindi: $f_Z = \frac{1}{|4\sqrt{\frac{z}{2}}|}\left[f_Y(\sqrt{\frac{z}{2}}) + f_Y(-\sqrt{\frac{z}{2}})\right]$





FUNZIONI DI UNA V.A.

- Data una variabile aleatoria Y ed una trasformazione g(y) formiamo Z=g(Y).
- La media di Z è $\mu_Z = E(Z) = \int_{-\infty}^{+\infty} z\, f_Z(z)\, dz$
- Se è nota la distribuzione di Y non è necessario conoscere la f_Z per determinare la media di Z:

$$\mu_Z = E\left[g(Y)\right] = \int_{-\infty}^{+\infty} g(y)\, f_Y(y)\, dy$$

- Invece il seguente problema è meno ovvio: **Non si sa di che tipo di variabile aleatoria sia Y, ma ne conosciamo media e varianza, è possibile determinare la media e la varianza di Z**?

FUNZIONI DI UNA V.A.

$$\mu_Z = E(Z) = \int_{-\infty}^{+\infty} z\, f_Z(z)dz$$

- IPOTESI: g(y) è lineare ovvero g(y)=aY+b
- Anche se non è nota la distribuzione di Y possiamo determinare la media di Z:

$$\mu_Z = E(Z) = E(aY+b) = \boxed{\int_{-\infty}^{+\infty} ay\, f_Y(y)dy + \int_{-\infty}^{+\infty} b\, f_Y(y)dy} = a\mu_Y + b$$

- Analogamente per la varianza

$$\sigma_z^2 = E\left[Z^2\right] - \mu_Z{}^2 = E\left[(aY+b)^2\right] - (a\mu_X + b)^2 =$$

$$= a^2 E\left[Y^2\right] - a^2\mu_X{}^2 + E\left[2abY\right] - 2ab\mu_X = a^2\sigma_Y^2$$

Importante

FUNZIONI DI UNA V.A.

- Quindi se la trasformazione è lineare è sufficiente conoscere media e varianza per determinare la media e la varianza della trasformazione della variabile aleatoria.

- Invece se la trasformazione non è lineare si possono solo stimare la media e la varianza di Z

- Si linearizza g intorno a μ_Y:

$$g(y) \approx g(\mu_Y) + \left.\frac{dg}{dy}\right|_{\mu_Y} (y-\mu_Y) + \frac{1}{2}\left.\frac{d^2g}{dy^2}\right|_{\mu_Y} (y-\mu_Y)^2$$

- Troncando al primo ordine:

$$\mu_Z = \int_{-\infty}^{\infty} g(y) f_Y(y) dy \cong \int_{-\infty}^{\infty} \left[g(\mu_Y) + \left.\frac{dg}{dy}\right|_{\mu_Y} (y-\mu_Y) \right] f_Y(y) dy = g(\mu_Y)$$

FUNZIONI DI UNA V.A.

- Troncando al secondo ordine: $\mu_Z \cong g(\mu_Y) + \frac{1}{2} \left. \frac{d^2 g}{dy^2} \right|_{\mu_Y} \sigma_Y^2$

- Per la varianza: $\sigma_Z^2 \cong \left(\left. \frac{dg}{dy} \right|_{\mu_Y} \right)^2 \sigma_Y^2$

TIPI DI V.A.

- TRASFORMAZIONI DI UNA GAUSSIANA
- Se Y è Gaussiana con media μ_Y e varianza σ_Y^2 allora la VA Z data dalla trasformazione Z=aY+b con a diverso da zero è anche essa Gaussiana con media a μ_Y +b e varianza $a^2\ \sigma_Y^2$.
- Da quanto detto, per la trasformazione Z=aY+b si ricava:

$$Y = \frac{Z-b}{a}$$

$$a = \frac{\sigma_Z}{\sigma_Y}$$

$$b = \mu_Z - \frac{\sigma_Z}{\sigma_Y}\mu_Y$$

$$\frac{Y-\mu_Y}{\sigma_Y} = \frac{Z-\mu_Z}{\sigma_Z}$$

Quindi nota una Gaussiana è possibile ricostruire qualunque altra Gaussiana. Per questo motivo si introduce la **Gaussiana Standard**

TIPI DI V.A.

- GAUSSIANA IN FORMA STANDARD
- Una gaussiana con media zero e varianza unitaria si dice Standard. Solitamente sono tabellati i valori di F e f per una Gaussiana standard

$$Y = N\left(\mu_Y, \sigma_Y^2\right), \quad Z = N(0,1)$$

$$F\left(\frac{Y-\mu_Y}{\sigma_Y}\right) = F(Z)$$

FUNZIONI DI UNA V.A. **IMPORTANTE**

- **Se la trasformazione è lineare** è sufficiente conoscere media e varianza per determinare la media e la varianza della trasformazione della variabile aleatoria.

$$\mu_Z = E(Z) = E(aY+b) = \int_{-\infty}^{+\infty} ay\, f_Y(y)dy + \int_{-\infty}^{+\infty} b\, f_Y(y)dy = a\mu_Y + b$$

$$\sigma_z^2 = E\left[Z^2\right] - \mu_Z{}^2 = E\left[(aY+b)^2\right] - (a\mu_X + b)^2 = a^2 E\left[Y^2\right] - a^2\mu_X{}^2 + E\left[2abY\right] - 2ab\mu_X = a^2\sigma_Y^2$$

- **Se la trasformazione non è lineare** si possono solo ottenere approssimazioni per la media e la varianza di Z

$$\mu_Z = \int_{-\infty}^{\infty} g(y) f_Y(y)dy \cong \int_{-\infty}^{\infty} \left[g(\mu_Y) + \left.\frac{dg}{dy}\right|_{\mu_Y} (y - \mu_Y) \right] f_Y(y)dy = g(\mu_Y)$$

$$\sigma_Z^2 \cong \left(\left.\frac{dg}{dy}\right|_{\mu_Y} \right)^2 \sigma_Y^2$$

- Una VA è standardizzata se ad essa è sottratta la sua media ed il risultato è diviso per la sua deviazione standard.

Variabile aleatoria χ^2

- Consideriamo n variabili aleatorie gaussiane **indipendenti** in forma standard $(X_1, X_2, \ldots, X_n)$
- La variabile aleatoria scalare

$$Z = X_1^2 + X_2^2 + \ldots + X_n^2$$

si chiama VA χ^2 con n gradi di libertà (spesso indicata simbolicamente con χ_n^2) ed ha pdf del tipo rappresentato nella diapositiva seguente.

- Per tale variabile la pdf **ha come solo parametro** n, gradi di libertà.

Storia da wikipedia

Karl Abbe (Eisenach, 23 gennaio 1840 – Jena, 14 gennaio 1905) è stato un ottico, fisico e astronomo tedesco.
Nato in una famiglia di umili origini, dal 1865 al 1896 fu professore di fisica all'Università di Jena e, dal 1878, fu anche direttore dell'osservatorio astronomico della città tedesca.

I suoi studi, rivolti prevalentemente all'ottica, gli fecero conoscere nel 1866 Carl Zeiss, a cui si associò nel 1875, assumendo la dirigenza della società di vetri ottici (sia per filtri sia per lenti) alla morte di quest'ultimo. Una volta rimasto di fatto l'unico padrone della società decise di rivoluzionarne la struttura organizzativa trasformandola in una cooperativa, che aveva tra i suoi membri, oltre che ai lavoratori, anche enti locali della zona e l'università di Jena.

A lui si deve anche la scoperta della variabile casuale Chi Quadrato (χ^2) analizzando la sommatoria di v.c. Normali standardizzate e indipendenti, che produce una nuova variabile causale, la χ^2 appunto.

Variabile aleatoria χ^2

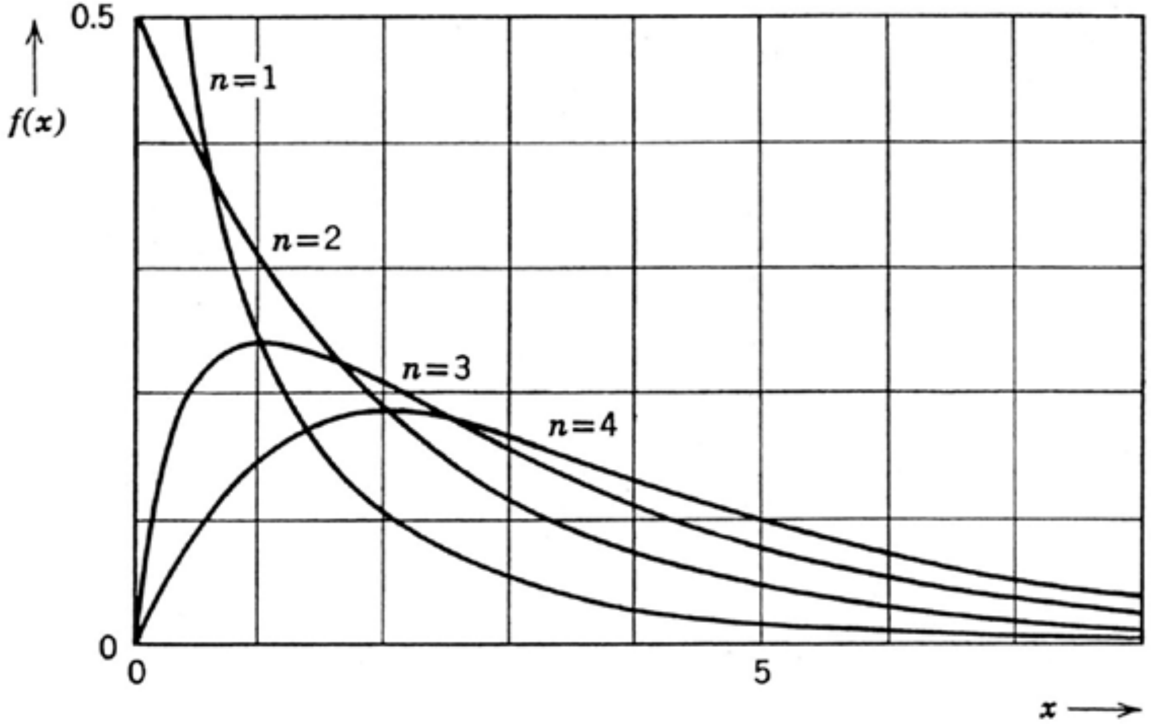


Fig. 10.1.1. Density of the chi-square distribution

Variabile aleatoria χ^2

- Proprietà di una VA χ^2 ad n gradi di libertà

$$\mu = n$$

$$\sigma^2 = 2n$$

$$f = \begin{cases} K_n x^{\frac{n-2}{2}} e^{-\frac{x}{2}} & x \geq 0 \\ 0 & x < 0 \end{cases}$$

$$Il\ massimo\ si\ ha\ per\ x = n-2,\ se\ n \geq 2$$

$$K_n = \frac{1}{2^{\frac{n}{2}}\Gamma(n/2)} \qquad \Gamma(\lambda) = \int_0^\infty e^{-t} t^{\lambda-1} dt$$

Variabile aleatoria t di Student

- Se la VA scalare X è gaussiana del tipo X ~ N(0, 1), e quindi ha media 0 e varianza unitaria ed Y è una VA scalare di tipo χ^2 ad n gradi di libertà e se, infine, X ed Y sono indipendenti, la variabile aleatoria scalare

$$Z = \frac{X}{\sqrt{Y/n}}$$

  si chiama VA **t di Student** con **n gradi di libertà** ed ha pdf del tipo rappresentato nella diapositiva seguente.

- La pdf è una funzione ad un parametro (n, gradi di libertà) ed è **simmetrica** rispetto allo 0.

- Ha media nulla $\mu = 0$ e varianza $\sigma^2 = n/(n-2)$ (per n>2 !)

Storia da wikipedia
William Sealy Gosset (Canterbury, 13 giugno 1876 – Londra, 16 ottobre 1937) è stato uno statistico inglese, meglio noto in statistica come lo Student del t di Student.
Gosset studia chimica e matematica al New College di Oxford.
Diversamente dagli altri suoi colleghi statistici famosi, Gosset non intraprende la carriera accademica, ma lavora presso la celebre birreria Guinness, dove elabora la mole di dati disponibili. Conclude la sua carriera gestendo dal 1935 la nuova birreria Guinness di Londra.
Presso Guinness si rende presto conto che le condizioni con le quali vengono raccolti i dati (temperatura, umidità, origine del malto,...) cambiano di continuo e il fatto di averne pochi con le stesse condizioni sperimentali non consentono di applicare il teorema del limite centrale che permette di far riferimento alla distribuzione gaussiana nei vari test statistici.
Nel 1908 pubblica con lo pseudonimo Student (passato alla storia della statistica) il lavoro in cui presenta la VA t

Variabile aleatoria t di Student

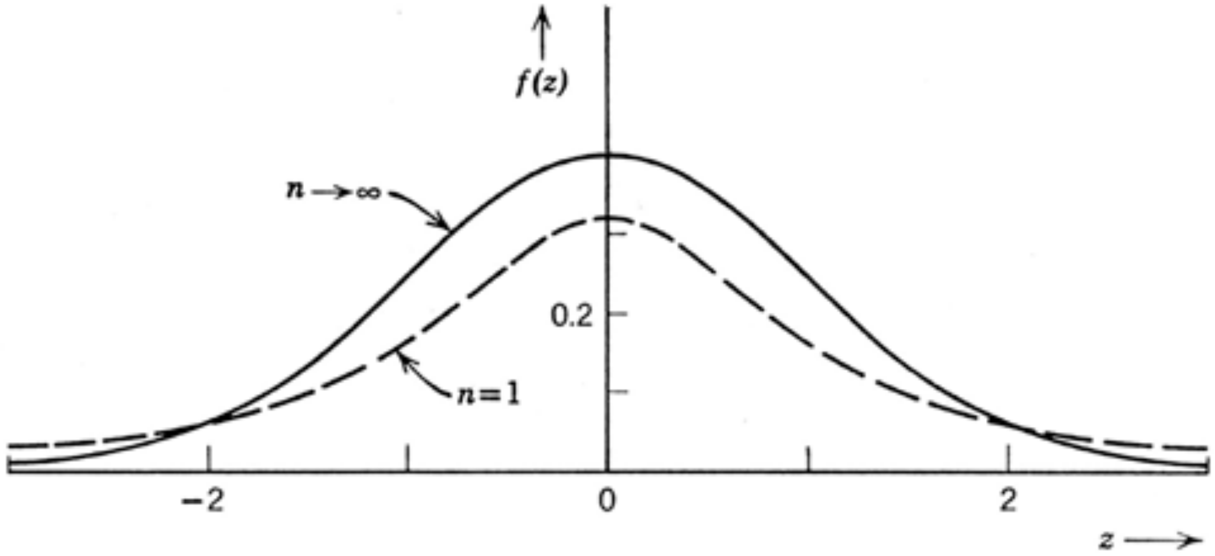


Fig. 10.3.1. Density of the t-distribution

Variabile aleatoria F di Fisher

- Se le due VA scalari Y e W sono VA di tipo χ^2 rispettivamente con m ed n gradi di libertà e sono indipendenti, la VA scalare Z

$$Z = \frac{Y/m}{W/n}$$

  è una VA di tipo **F di Fisher** con (m, n) gradi di libertà.

- La pdf ha due parametri: m ed n.

- La media di una VA F con (m, n) gradi di libertà è $\mu = n/(n-2)$, per $n > 2$!

- Tipici grafici della pdf delle VA F per vari gradi di libertà sono riportati nella diapositiva seguente.

Storia da wikipedia
Ronald Aylmer Fisher (Londra, 17 febbraio 1890 – Adelaide, 29 luglio 1962) è stato uno statistico e matematico inglese.
Viene considerato colui che ha fatto della statistica una scienza moderna, in quanto ha fondato i concetti di riferimento della statistica matematica moderna.
È tra i primi, o il primo, a comprendere l'importanza del campionamento casuale per poter generalizzare i risultati, in opposizione ai campionamenti fatti secondo criteri vari di opportunità.
Nel 1925 perfeziona il metodo ideato da William Sealy Gosset (alias Student) per confrontare due medie, ideando il test "t di Student" attualmente usato e introducendo il concetto di gradi di libertà.
Importante innovazione è la cosiddetta analisi della varianza, ma è un suo allievo (George W. Snedecor) a utilizzare una distribuzione diversa da quella gaussiana, introducendo la variabile casuale F di Snedecor, dove la F è in onore al maestro (Fisher).

Variabile aleatoria F di Fisher

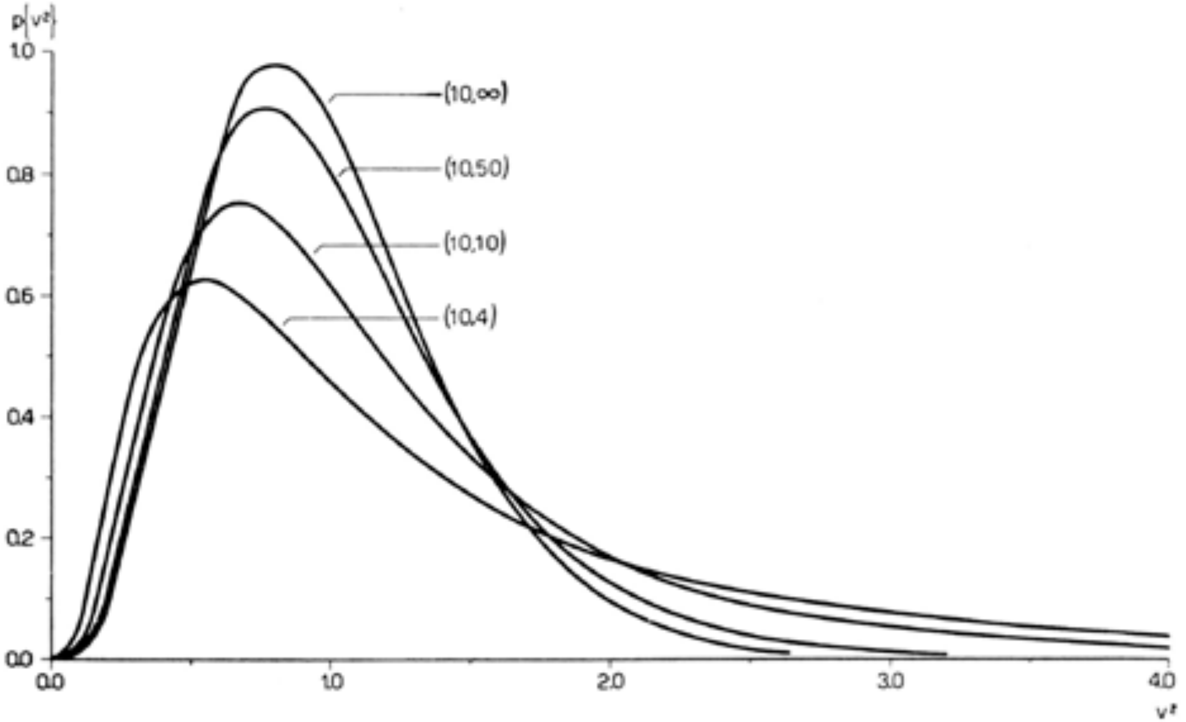


Fig. 14.1. Distribution curves of v^2 for (f_1, f_2) = (10, 4), (10, 10), (10, 50), and (10, ∞).

Concetti importanti

- Modello dell'esperimento
- Teoria della probabilità
- Variabile aleatoria
- Distribuzioni di probabilità
  - Proprietà delle popolazioni
- Tipi di variabili aleatorie
- Trasformazioni di Variabili aleatorie
- Trasformazioni di VA
  - Lineari: preservano il tipo di VA
  - Non lineari: cambiano il tipo di VA
- Caratterizzazione sintetica e trasformazione.
  - Gaussiane: la conoscenza di media e varianza permette la trasformazione esatta
  - Non gaussiane: si perviene ad una conoscenza approssimata dei parametri media e varianza.
- Variabili aleatorie di tipo chi quadro, t di student e F di Fisher

Appendice comandi Matlab

- Con Matlab (Toolbox Statistics) è possibile rappresentare in maniera semplice le funzioni densità di probabilità e distribuzioni cumulative.
- Se si eseguono per esempio le seguenti istruzioni:

```
>> x = [-3:0.02:3];
>> y = normpdf(x);
>> plot(x,y)
```

  - Si crea un vettore riga x di valori che partono da -3 sino a 3, a intervalli di 0.02
  - Si calcola un vettore y che associa ad ogni elemento di x il valore della distribuzione normale di media 0 e varianza 1.
  - diagramma infine y contro x.
- Il comando

```
>> y = normpdf(x, mu, sigma)
```

  - Restituisce la distribuzione densità di probabilità di tipo normale con media mu e deviazione standard, sigma, ai valori corrispondenti in x.
  - Il vettore y ha le stesse dimensioni di x.

Appendice comandi Matlab

- Il comando

```
>> y = chi2pdf(x, v)
```

  - Restituisce la distribuzione densità di probabilità di tipo chiquadro a v gradi di libertà

- Il comando

```
>> y = tpdf(x, v)
```

  - Restituisce la distribuzione densità di probabilità di tipo T-student a V gradi di libertà.

- Il comando

```
>> y = fpdf(x, v1, v2)
```

  - Restituisce la distribuzione densità di probabilità di tipo F a V1 e V2 gradi di libertà

Appendice comandi Matlab

- In modo assolutamente analogo è possibile calcolare i valori e rappresentare graficamente le distribuzioni cumulative.
- I comandi

```
>> y = normcdf(x, mu, sigma)
>> y = chi2cdf(x, v)
>> y = tcdf(x, v)
>> y = fcdf(x, v1, v2)
```

  - Restituiscono rispettivamente le cumulative per le distribuzioni gaussiane, le chi-quadro, le T-student, e le distribuzioni di Fisher

Appendice comandi Matlab

- È possibile anche calcolare le funzioni inverse delle cdf, ovvero, data una certa probabilità α, quale è il numero c tale che $Pr(y \leq c) = \alpha$?

- Il comando per ricavare l'inversa della distribuzione cumulativa, nel caso di distribuzioni normali è:

```
>> norminv(α, mu, sigma)
```

  – dove α è la probabilità che si intende valutare