# LEZIONE 8

# FUNZIONI DI UNA VARIABILE ALEATORIE

Nota la pdf di una variabile aleatoria X,
e data Y=g(X), Y e' un variabile aleatoria.

Qual e' il valore di aspettazione, qual e' la varianza di Y ?

# FUNZIONI DI VARIABILI ALEATORIE

Nota la pdf di una variabile aleatoria X,
e Y=g(X), Y e' un variabile aleatoria.

Qual e' il valore di aspettazione, qual e' la varianza di Y ?

$$E[Y] = E[g(x)] = \int y f_Y(y) dy$$

**law of the unconscious statistician**

LOTUS stabilisce che

$$E[Y] = E[g(x)] = \int y f_Y(y) dy = \int g(x) f_X(x) dx$$

===>>>> non occorre conoscere la pdf della variabile y per calcolare il valore di aspettazione di Y, basta calcolare il valore di aspettazione della g(x) secondo la pdf della variabile aleatoria X.

**DOMANDA:** ???

$E[Y] = E[\ g(x)\ ] = g(\ E[X]\ )$ ??? **NON SEMPRE**

# FUNZIONI DI VARIABILI ALEATORIE

**DOMANDA:** ???

$E[Y] = E[g(x)] = g(E[X])$ ??? **NON SEMPRE**

Osserviamolo empiricamente su un esempio

Esempio:
Supponiamo di chiedere ad una classe di ragazzini di 5 elementare di tagliarci dei dischi di carta di 10 centimetri di diametro.
Alla fine del lavoro constatiamo che **la variabile aleatoria raggio dei dischi** segue una distribuzione gaussiana con media 5 cm e sigma 1 cm.

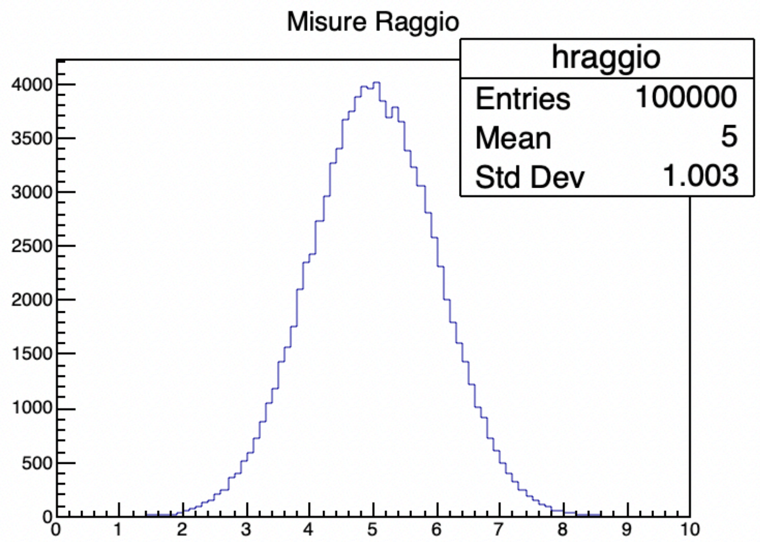
1) Che distribuzione di probabilità seguirà la variabile aleatoria superficie dei dischi?

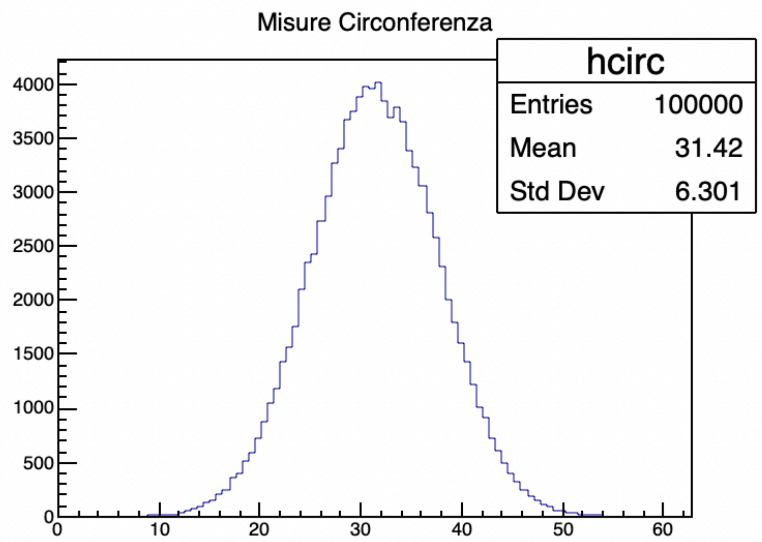
2) Quanto varranno la media e la varianza della variabile aleatoria superficie dei dischi

**Simulazione dell'esperimento**.
La variabile aleatoria che rappresenta il raggio dischi di carta può essere simulata estraendo N valori da una distribuzione gaussiana con media 5 e varianza 1.

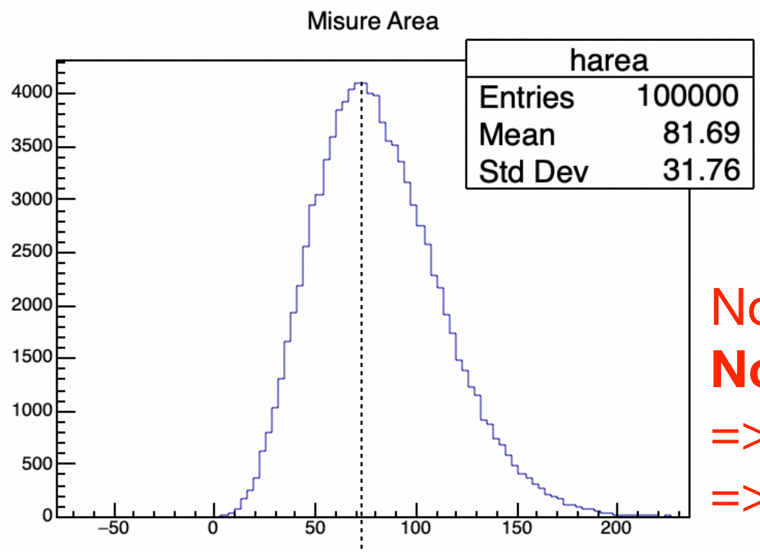
L'utilizzo della simulazione mi permette di determinare per il mio campione di dischi la loro superficie e quindi di studiare direttamene la densità di probabilità associata alla variabile aleatoria superficie del disco

# FUNZIONI DI VARIABILI ALEATORIE



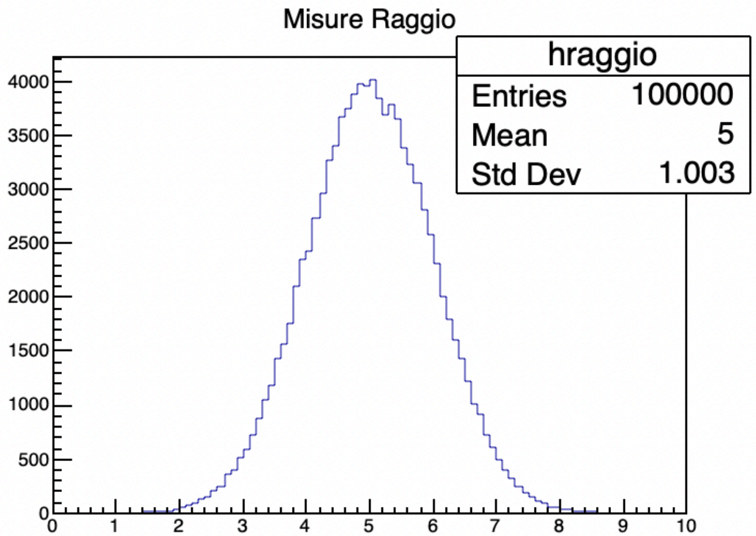



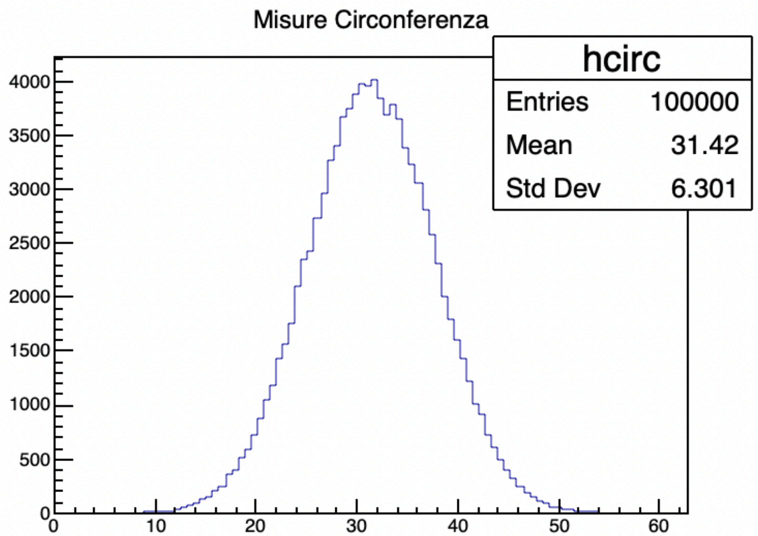

Nota: circonferenza = 2πR
E' distribuita gaussianamente come R
=> Media 31.42 = 2π<R>

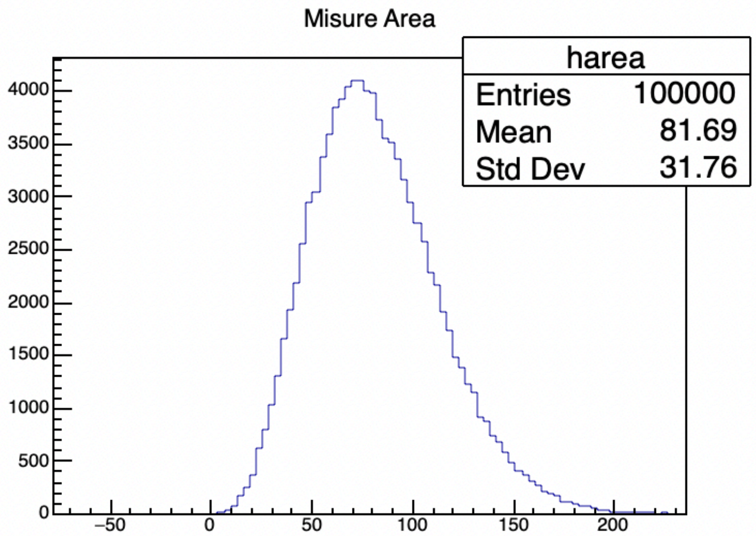



Nota: **area** = $\pi R^2$
**Non** e' distribuita gaussianamente come R
=> Media 81.69 > Moda ~ 72
=> Media != $\pi \langle R \rangle^2$ ~ 78,5

# FUNZIONI DI VARIABILI ALEATORIE




```
Raggio, raggioMedio      5 5.0002 devStd_raggio 1.00284

                Raggio,  Circ, Area        5      31.4159 78.5398 valori esatti
da simulazione: Raggio,  Circ, Area       5.0002 31.4172 81.7054
da simulazione:eRaggio, eCirc,eArea       0.00317126 0.0993498 0.258375
da f(R_medio) :          Circ, Area       5.0002 31.4172 78.5459
simul-f(<R>)  :          Circ, Area       2.45137e-13 3.15946
[simul-f(<R>)]/err:      Circ, Area       2.46742e-12 12.2282
Correzione per Area = 3.15946   Area media attesa(post correzione) = 81.7054
```

# FUNZIONI DI VARIABILI ALEATORIE

Nota la pdf di una variabile aleatoria X,
e Y=g(X), Y e' un variabile aleatoria.

Qual e' il valore di aspettazione, qual e' la varianza di Y ?

$$E[Y] = E[g(x)] = \int y f_Y(y) dy$$

**law of the unconscious statistician**

***LOTUS*** stabilisce che

$$E[Y] = E[g(x)] = \int y f_Y(y) dy = \int g(x) f_X(x) dx$$ ===>>>> non occorre conoscere la pdf della variabile y per calcolare il valore di aspettazione di Y, basta calcolare il valore di aspettazione della g(x) secondo la pdf della variabile aleatoria X.

# FUNZIONI DI VARIABILI ALEATORIE

LOTUS Dimostrazione non rigorosa

$$E[Y] = E[g(x)] = \int y f_Y(y)dy = \int g(x) f_X(x) dx$$

Variabile Y=g(X), dove la g(X) è una funzione **strettamente monotona crescente**.
Supponiamo nota la densità di probabilità che descrive la variabile X ($f_X(X)$) che espressione avrà la distribuzione di probabilità associata alla Y ($f_Y(Y)$)?

Ricaviamo $f_Y(y)$ con una dimostrazione non rigorosa e limitata al caso di una g(x) strettamente monotona e pertanto invertibile

Parto dall'identità:

$$\int_{-\infty}^{+\infty} f_X(x)dx = 1 = \int_{-\infty}^{+\infty} f_Y(y)dy$$

Essendo y=g(x) strettamente monotona allora è invertibile per cui posso determinare $x=g^{-1}(y)$, da cui:

$$dx = \frac{d(g^{-1}(y))}{dy} dy$$

Eseguendo il cambiamento di variabili al primo membro della precedente identità ottengo

$$\int_{-\infty}^{+\infty} f_X(g^{-1}(y)) \frac{d(g^{-1}(y))}{dy} dy = \int_{-\infty}^{+\infty} f_Y(y)dy$$

Non rigoroso….

$$f_Y(y) = f_X(g^{-1}(y)) \frac{d(g^{-1}(y))}{dy}$$

Dipartimento di Matematica e Fisica "Ennio De Giorgi"

INFN

LOTUS Dimostrazione non rigorosa

$$E[Y] = E[g(x)] = \int y f_Y(y) dy = \int g(x) f_X(x) dx$$

Se Y=g(X), con g(X) **strettamente monotona decrescente**.

$$\int_{-\infty}^{+\infty} f_X(g^{-1}(y)) \frac{d(g^{-1}(y))}{dy} dy = -\int_{-\infty}^{+\infty} f_Y(y) dy$$

$$f_Y(y) = -f_X(g^{-1}(y)) \frac{d(g^{-1}(y))}{dy}$$

Da cui genericamente:

$$f_Y(y) = f_X(g^{-1}(y)) \left| \frac{d(g^{-1}(y))}{dy} \right|$$

# FUNZIONI DI VARIABILI ALEATORIE

LOTUS Dimostrazione non rigorosa

$$E[Y] = E[g(x)] = \int y f_Y(y) dy = \int g(x) f_X(x) dx$$

Calcolo della media della variabile aleatoria Y=g(X)

Per calcolare il valore di aspettazione della variabile aleatoria Y occorre fare l'integrale

$$\mu_y = \int_{-\infty}^{+\infty} y f_Y(y) dy$$

ma

$$f_Y(y) = f_X(g^{-1}(y)) \left| \frac{d(g^{-1}(y))}{dy} \right|$$

Relazione valida sempre anche se g(x) non e' strettamente monotona

$$x = g^{-1}(y); \qquad dx = \left| \frac{d(g^{-1}(y))}{dy} \right| dy \quad \Rightarrow \quad \boxed{\mu_y = \int_{-\infty}^{+\infty} g(x) f_X(x) dx}$$

L'ultima espressione mi permette di determinare, quindi, il valore di aspettazione della
variabile Y=g(X) nota la densità di probabilità che caratterizza la variabile aleatoria X.

# FUNZIONI DI VARIABILI ALEATORIE

Calcolo della varianza della variabile aleatorie Y=g(X)

Analogamente al caso della media, per calcolare la varianza della Y dovrei calcolare l'integrale

$$\mathrm{var}(Y) = \int_{-\infty}^{+\infty} (y-\mu_y)^2 f_Y(y)dy$$

ma

$$f_Y(y) = f_X(g^{-1}(y)) \left| \frac{d(g^{-1}(y))}{dy} \right|$$

$$x = g^{-1}(y) \qquad ; \qquad dx = \left| \frac{d(g^{-1}(y))}{dy} \right| dy$$

quindi

$$\mathrm{var}(Y) = \int_{-\infty}^{+\infty} (g(x)-\mu_y)^2 f_X(x)dx$$

Relazione valida sempre anche se g(x) non e' strettamente monotona

Quest'ultima espressione mi permette di determinare la varianza utilizzando la densità di probabilità della variabile X

# FUNZIONI DI VARIABILI ALEATORIE

Media e varianza di una funzione lineare di X

**Alcuni casi particolari**

Per valor medio e varianza valgono le seguenti proprietà:

$$Y = aX + b$$

$$E[Y] = E[aX + b] = aE[X] + b \quad \forall a, b \in \Re$$

**DOMANDA:**
$E[Y] = E[g(x)] = g(E[X])$ ?
**SI per funzioni lineari**

Dimostrabile dalla definizione di valore di aspettazione.
Qualche attenzione in più merita il calcolo della varianza.

$$\mathrm{var}(Y) = \mathrm{var}(aX + b) = E[(aX + b - E[aX + b])^2] =$$

$$= E[(aX + b - aE[X] - b)^2] =$$

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

$$= a^2\, \mathrm{var}(X)$$

Per cui

$$\mathrm{var}(aX + b) = a^2\, \mathrm{var}(X) \quad \forall a, b \in \Re$$

# FUNZIONI DI VARIABILI ALEATORIE

Media e varianza di una funzione ***generica*** di X

**Caso generale**

$$\mu_y = \int_{-\infty}^{+\infty} g(x) f_X(x) dx$$

Ipotizziamo ora che la variabile aleatoria X sia caratterizzata da una densità di probabilità di tipo gaussiano con standard deviation in valore assoluto piccola rispetto al valore medio. Per calcolare l'integrale precedente possiamo sviluppare in serie la funzione g(x) intorno alla media della variabile aleatoria X

$$g(x) = g(\mu_X) + (x - \mu_X) g'(\mu_X) + \frac{1}{2}(x - \mu_X)^2 g''(\mu_X) + \ldots$$

$$\mu_Y = \int g(\mu_X) f(x) dx + \int (x - \mu_X) g'(\mu_X) f(x) dx + \int \frac{1}{2}(x - \mu_X)^2 g''(\mu_X) f(x) dx + \ldots$$

Se mi fermo ai primi due termini dello sviluppo in serie ottengo che

$g(\mu_X) \int f(x) dx = g(\mu_X)$ Termine di ordine 0

$g'(\mu_X) \int (x - \mu_X) f(x) dx = 0$ Termine di ordine 1

Per cui posso concludere che:

$$\mu_Y = g(\mu_X)$$

Se mi fermo al primo ordine

# FUNZIONI DI VARIABILI ALEATORIE

Ma attenzione. Non sempre i termini di ordine superiore possono essere trascurati!
Se aggiungessi il termine del secondo ordine dello sviluppo precedente

$$\frac{1}{2}\int (x-\mu_X)^2 g''(\mu_X) f(x)dx = \frac{1}{2} g''(\mu_X) \int (x-\mu_X)^2 f(x)dx =$$

$$= \frac{1}{2} g''(\mu_X)\,\mathrm{var}(X)$$

Termine di ordine 2

Otterrei che

$$\mu_Y = g(\mu_X) + \frac{1}{2} g''(\mu_X)\,\mathrm{var}(x)$$

**DOMANDA:**
$E[Y] = E[g(x)] = g(E[X])$ ?
**NO in generale**

Questo termine aggiuntivo risulta trascurabile solo se il prodotto $g''(\mu_X)\mathrm{var}(X)$ è piccolo rispetto a $g(\mu_X)$. Quindi solo nel caso di funzioni g(x) lentamente variabili ($g''(\mu_X) \cong 0$) nella regione di interesse posso trascurare il termine correttivo.

# FUNZIONI DI VARIABILI ALEATORIE

- **Varianza**
  - Come prima sviluppiamo in serie di Taylor g(x) attorno a $\mu_X$
    - $$g(x) = g(\mu_X) + (x-\mu_X)g'(\mu_X) + \frac{1}{2}(x-\mu_X)^2 g''(\mu_X) + \ldots$$
  - *var(Y)* = $\int [\ g(\mu_X) - \mu_Y + (x-\mu_X)g'(\mu_X) + \frac{1}{2}(x-\mu_X)^2 g''(\mu_X) + \ldots\ ]^2\ f(x)dx$
    - ***ci arrestiamo al primo ordine***, quindi per consistenza usiamo $\mu_Y = g(\mu_X)$ quindi
  - $= \int [\ g(\mu_X) - \mu_Y\ ]^2 f(x)dx + \int [(x-\mu_X)g'(\mu_X)]^2 f(x)dx + \int 2[\ g(\mu_X) - \mu_Y\ ][(x-\mu_X)g'(\mu_X)]\ f(x)dx$
    - Il primo termine e' nullo perché $g(\mu_X) - \mu_Y$ e' circa zero
    - Il terzo e e' nullo perché si tratta di un momento centrato del primo ordine, inoltre ricompare $g(\mu_X) - \mu_Y$

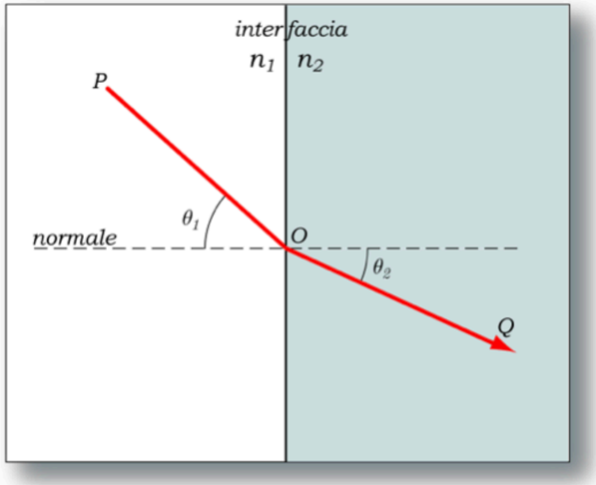
$$\mathrm{var}(Y) = g'^2(\mu_X)\,\mathrm{var}(X)$$

**Formula "standard" alla base della propagazione degli errori statistici.**
**… ma come è facile immaginare la cosa non funziona perfettamente …..**

Perche' e' ottenuta con uno sviluppo fino al primo ordine

# FUNZIONI DI VARIABILI ALEATORIE

Esempio: La legge di Snell
La legge di Snell descrive le modalità di rifrazione di un raggio luminoso quando passa tra due mezzi di indice di rifrazione diverso




$$n_1 \sin(\theta_1) = n_2 \sin(\theta_2)$$

Dove $n_1$ e $n_2$ sono detti indici di rifrazione rispettivamente del mezzo 1 e 2.
Se il mezzo 1 è il vuoto allora l'indice di rifrazione è pari ad 1. (Nel caso dell'aria l'indice di rifrazione è quasi 1 e spesso viene approssimato con 1)

Supponiamo di avere uno strumento che mi permette di misurare l'indice di rifrazione di un mezzo facendo incidere un raggio luminoso in modo quasi radente sul mezzo.
Ipotizziamo che lo strumento misuri l'angolo di incidenza con un errore relativo del 5%, mentre misuri l'angolo di rifrazione con precisione molto maggiore.
Possiamo ritenere che al calcolo dell'errore su n contribuisca solo la misura dell'angolo di incidenza.

# FUNZIONI DI VARIABILI ALEATORIE

Supponiamo che si effettui una serie di misure e si ottenga che la variabile aleatoria *angolo di incidenza* segua una distribuzione di tipo gaussiano con media pari a 1.56 radianti e deviazione standard di 0.08 radianti.
Se l'angolo di rifrazione è pari a 0.6283185 radianti (con errore trascurabile) quanto vale l'indice di rifrazione e qual è l'errore associato a questa misura?

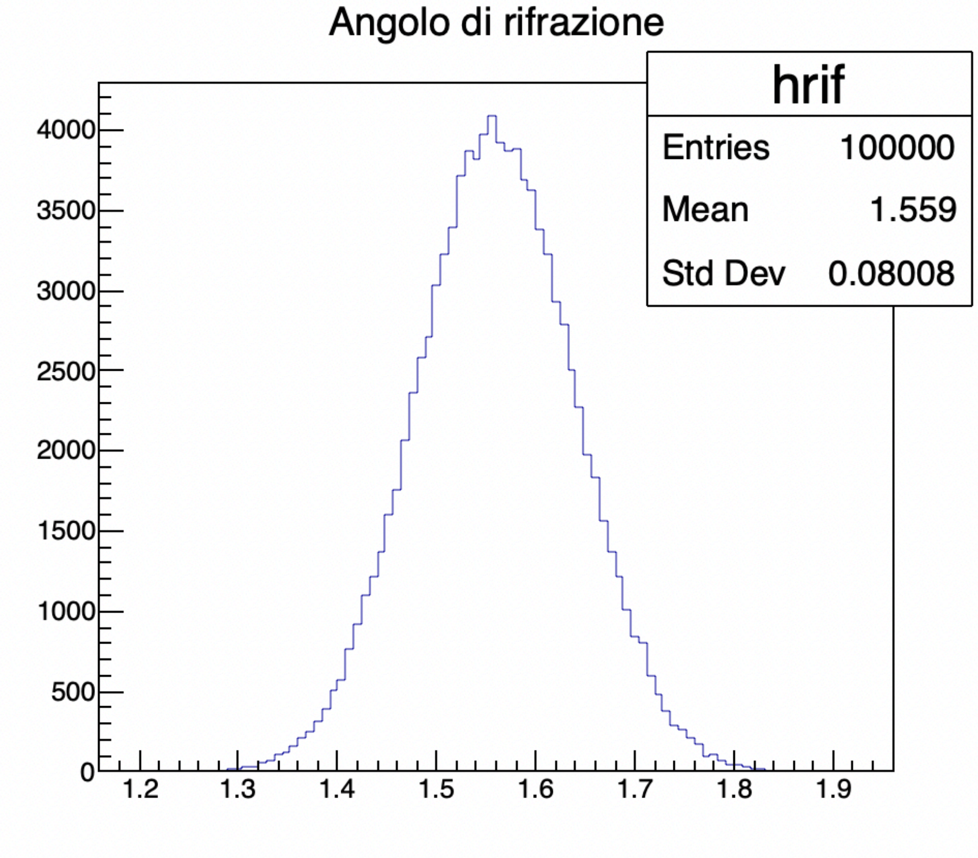
Calcolo analitico con formule approssimate:

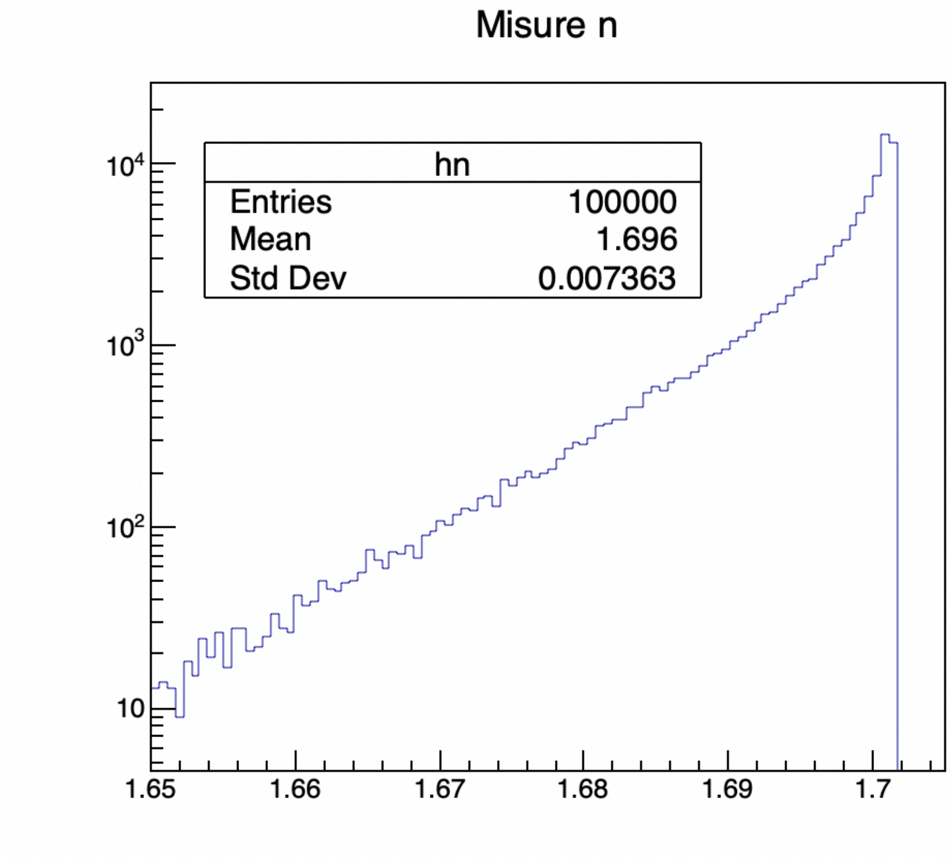
$$n = \frac{\sin(\theta_i)}{\sin(\theta_r)} = 1.7012$$

$$\mathrm{var}(n) = \left(\frac{\cos(\theta_i)}{\sin(\theta_r)}\right)^2 \mathrm{var}(X) = 0.000002$$

Simuliamo l'esperimento per valutare correttamente la propagazione degli errori ….

# FUNZIONI DI VARIABILI ALEATORIE




Risultati attesi: n=1.70119 errore=4.90223e-06
Risultati ottenuti con la simulazione:  indice di rifrazione ed errore 1.69575 2.47642e-05
Correction (for second order term)= -0.00545502
Difference Simulation - Theory(1st order) = 0.0054463
Valore atteso + correzione = 1.69574

# FUNZIONI DI PIU' VARIABILI ALEATORIE

# FUNZIONI DI DUE VARIABILI ALEATORIE

## Media

Consideriamo il caso di una variabile aleatoria derivata funzione di due variabili aleatorie z=g(x,y), dove g(x,y) è una funzione qualunque.
Si può dimostrare che calcolare

$$\mu_Z = \int z f_Z(z) dz$$

Equivale a calcolare

$$\mu_Z = \iint g(x,y) f(x,y) dx dy$$

E ricorrendo allo sviluppo in serie di Taylor della funzione g(x,y) intorno al punto $P_\mu=(\mu_X, \mu_Y)$ fermandosi al primo ordine si ottiene.

$$\mu_Z = g(\mu_X, \mu_Y)$$

Dove $\mu_X$ e $\mu_Y$ sono i valori di aspettazione della distribuzioni marginali.

## Dimostrazione

$$\mu_X = \int x f_X(x)dx = \iint x f(x,y)dxdy$$ Valore di aspettazione della variabile X

$$\mu_Y = \int y f_Y(y)dy = \iint y f(x,y)dxdy$$ Valore di aspettazione della variabile Y

$$g(x,y) = g(\mu_X,\mu_Y) + (x-\mu_X)\left.\frac{\partial g}{\partial x}\right|_{\mu_X,\mu_Y} + (y-\mu_Y)\left.\frac{\partial g}{\partial y}\right|_{\mu_X,\mu_Y} + \boxed{?}$$

$$\mu_Z = \iint g(x,y)f(x,y)dxdy = g(\mu_X,\mu_Y)\underbrace{\iint f(x,y)dxdy}_{1} +$$

$$\iint \underbrace{(x-\mu_X)f(x,y)}_{0}\left.\frac{\partial g}{\partial x}\right|_{\mu_X,\mu_Y} dxdy + \iint \underbrace{(y-\mu_Y)f(x,y)}_{0}\left.\frac{\partial g}{\partial y}\right|_{\mu_X,\mu_Y} dxdy + \boxed{?}$$

$$\mu_Z = g(\mu_X,\mu_Y)$$

Ma in questo modo, come nel caso ad una singola variabile, è possibile introdurre sistematiche se le derivare seconde della g(x,y) nel punto $P_\mu$ sono molto diverse da zero.

# FUNZIONI DI DUE VARIABILI ALEATORIE

## Casi semplici: media e varianza di funzioni lineari

$$E[aX+bY]=aE[X]+bE[Y] \quad \forall a,b \in \Re$$

Dimostrabile dalla definizione di valore di aspettazione

Uso E[X+Y] = E[X]+E[Y]

$$\boxed{\mathrm{var}(X+Y)} = E[((X+Y)-E[X+Y]))^2] =$$

$$= E[((X+Y)-E[X]-E[Y])^2] = E[\,(\,(X-E[X]) + (Y-E[Y])\,)^2\,] =$$

$$= E[((X-E[X])^2 + (Y-E[Y])^2 + 2(X-E[X])(Y-E[Y])] =$$

$$= \boxed{\mathrm{var}(X)+\mathrm{var}(Y)+2\,\mathrm{cov}(X,Y)}$$

$$\mathrm{cov}(X,Y) = \int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty}(y-\mu_y)(x-\mu_x)f(x,y)\,dx\,dy$$

Covarianza

Dove f(x,y) è la densità di probabilità congiunta delle due variabili aleatorie e $\mu_x$ $\mu_y$ rappresentano rispettivamente il valore di aspettazione della variabile X e della variabile Y.

## Covarianza

$\boxed{\text{Se le variabili sono indipendenti}}$ allora

$$f(x,y) = f_X(x) f_Y(y)$$

Da cui segue che la covarianza è nulla essendo il valore di aspettazione intorno alla media di una qualunque variabile aleatoria nullo.

$$E[(x-\mu)] = 0$$

$$\mathrm{cov}(X,Y) = \int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty}(y-\mu_y)(x-\mu_x) f(x,y)\,dx\,dy =$$

$$= \int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty}(y-\mu_y)(x-\mu_x) f_X(x) f_Y(y)\,dx\,dy =$$

$$= \int_{-\infty}^{+\infty}(y-\mu_y) f_Y(y)\,dy \int_{-\infty}^{+\infty}(x-\mu) f_X(x)\,dx = 0$$

# FUNZIONI DI DUE VARIABILI ALEATORIE

## Covarianza

Nel caso di variabili aleatorie indipendenti

$$E[aX+b]=aE[X]+b \quad \forall a,b \in \Re$$

$$\mathrm{var}(aX+b)=a^2\,\mathrm{var}(X) \quad \forall a,b \in \Re$$

$$E[aX+bY]=aE[X]+bE[Y] \quad \forall a,b \in \Re$$

$$\mathrm{var}(aX+bY)=a^2\,\mathrm{var}(X)+b^2\,\mathrm{var}(Y) \quad \forall a,b \in \Re$$

Da cui consegue che:

$$E[X-Y]=E[X]-E[Y]$$

$$\mathrm{var}(X-Y)=\mathrm{var}(X)+\mathrm{var}(Y)$$

Utilizzato nella stima della propagazione degli errori sperimentali su grandezze derivate

**Si noti che la covarianza nulla è condizione necessaria ma non sufficiente affinché due variabili aleatorie siano indipendenti.**

## Covarianza e coefficiente di correlazione

Spesso invece della covarianza si usa il coefficiente di correlazione

$$\rho = \frac{\mathrm{cov}(x,y)}{\sqrt{\mathrm{var}(x)\,\mathrm{var}(y)}}$$

Se le variabili sono indipendenti la correlazione è nulla.
Il coefficiente di correlazione, data la sua definizione, assume valori compresi tra -1 e 1 ed è adimensionale (non lo sono media e varianza!)

# FUNZIONI DI DUE VARIABILI ALEATORIE

## Varianza di Z=g(X,Y) in generale

In termini più generali, limitandosi al primo ordine si trova che

$$\mathrm{var}(Z) = \left(\frac{\partial g}{\partial x}\right)^2_{|P_\mu} \mathrm{var}(X) + \left(\frac{\partial g}{\partial y}\right)^2_{|P_\mu} \mathrm{var}(Y) + 2\,\mathrm{cov}(x,y)\left(\frac{\partial g}{\partial x}\frac{\partial g}{\partial y}\right)_{|P_\mu}$$

Si noti la presenza del termine di covarianza già approssimando al primo ordine lo sviluppo in serie.
Se le variabili sono indipendenti allora il termine di covarianza scompare.

Nel caso più generale di $Z=g(X_1,X_2, \dots ,X_n)$ l'espressione della varianza assume la forma.

$$\mathrm{var}(Z) = \sum_{i,j} \frac{\partial f}{\partial x_i}\frac{\partial f}{\partial x_j}\mathrm{cov}(x_i, x_j)$$

Dove

$$\mathrm{cov}(x_i, x_i) = \mathrm{var}(x)$$

In generale per una stima accurata dei momenti delle variabili aleatorie derivate è conveniente ricorrere a tecniche di simulazione.

## Stima della Varianza e Covarianza da un campione

Per calcolare la covarianza tra due grandezze misurate consideriamo (per ora) il parallelismo con le stime per una singola variabile aleatoria.

**Stima del valore di aspettazione della variabile aleatoria associata al lancio di un dado.**

1. Per calcolare il valore di aspettazione (media) eseguo N lanci e calcolo la media.

   N=21; 1 + 4 + 4 + 6 + 2 + 3 + 2 + 5 + 1 + 4 + 6 + 5 + 3 + 1 + 6 + 5 + 2 + 3 + 6 + 1 + 2 = 72; 72/21=3.42

2. Per calcolare il valore di aspettazione calcolo le frequenze di uscita dei singoli valori negli N lanci.

   N=21; $f_1$ =4/21; $f_2$ =4/21; $f_3$ =3/21; $f_4$ =3/21; $f_5$ =3/21; $f_6$ =4/21; 1* $f_1$ + 2* $f_2$ + 3* $f_3$ + 4* $f_4$ + 5* $f_5$ + 6* $f_6$ = 3.42

3. Per calcolare il valore di aspettazione assumo nota la distribuzione di probabilità del processo.

$$\mu_x = \sum x_i p_i;$$

$$p_i = \frac{1}{6}$$

$$\mu_x = \frac{1}{6}(1+2+3+4+5+6) = 3.5$$

# Funzioni di due variabili aleatorie

## Stima della Varianza e Covarianza da un campione

Varianza di un campione

$$\mathrm{var}(x) = \int_{-\infty}^{+\infty} (x-\mu_x)^2 f(x)dx$$

$$\mathrm{var}(x) = \frac{\sum (x_i-\mu_x)^2}{N}$$

Covarianza tra due grandezze.

$$\mathrm{cov}(X,Y) = \int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty} (y-\mu_y)(x-\mu_x) f(x,y)dxdy$$

$$\mathrm{cov}(X,Y) = \frac{\sum (y_i-\mu_y)(x_i-\mu_x)}{N}$$

Si dimostra che cov(x,y)= $\frac{1}{N}\sum_{i=1}^{N} x_i y_i - \left(\frac{1}{N}\sum_{i=1}^{N} x_i\right)\left(\frac{1}{N}\sum_{i=1}^{N} y_i\right)$

# TEORIA DEGLI ERRORI DI MISURA

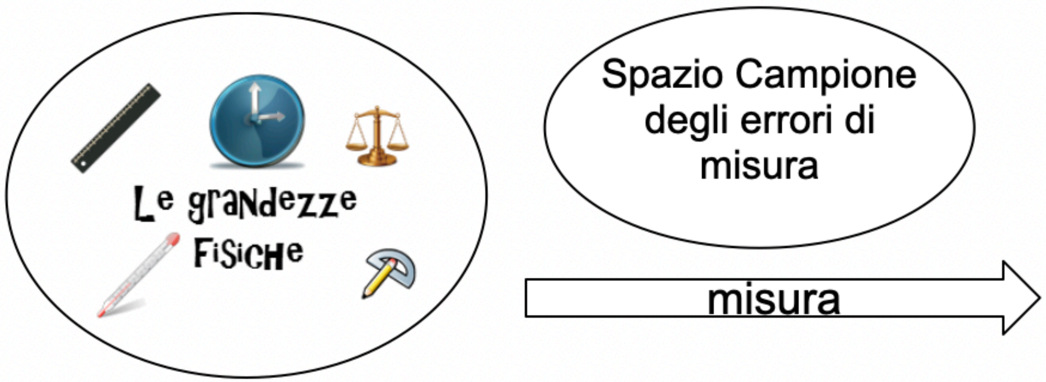
Nella teoria degli errori si assume che la misura
i-esima di una grandezza fisica il cui valore vero è μ, in conseguenza di indeterminabili errori di misura casuali dovuti al metodo di misura stesso e al processo fisico in studio, assuma il valore $x_i=\mu+\delta_i$, dove $\delta_i$ è l'errore associato alla misura i-esima.

*Quando eseguo una misura $x_i$ di una grandezza fisica con valore vero $\mu$ so che in conseguenza del processo di misura questa grandezza sarà affetta da un errore $\delta_i$ il quale è un elemento dello spazio campione degli errori le cui caratteristiche dipendono dal metodo di misura adottato.*

# TEORIA DEGLI ERRORI DI MISURA

*Quando dico che il valore misurato della grandezza fisica è $x_i \pm \sigma$ con $\sigma$ errore statistico sto facendo alcune assunzioni.*

*Sto assumendo che* ***lo spazio campione degli errori di misura segue una distribuzione di probabilità nota*** *e che in un qualche modo* ***sono in grado di conoscerne la sigma****. Sto inoltre implicitamente assumendo che tale spazio campione abbia* ***valore di aspettazione nullo****.*




$x_i \pm \sigma$

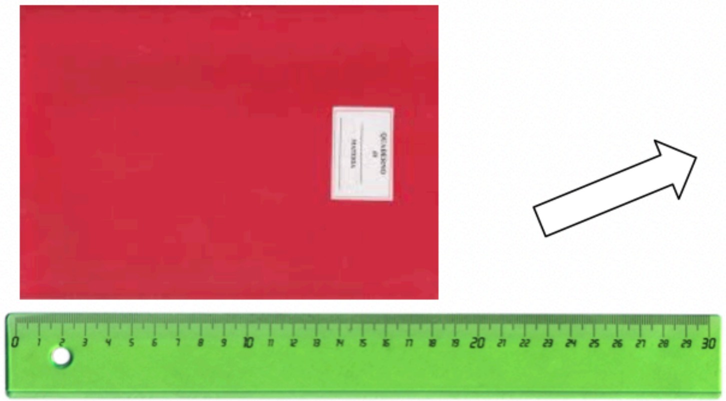
L'assumere che il valore di aspettazione della popolazione "*errori di misura*" associata al processo di misura in studio sia a valore di aspettazione nullo equivale a dire che la misura non sia affetta da "*errori sistematici*".

# TEORIA DEGLI ERRORI DI MISURA

Se la misura è unica significa che ho assunto la misura stessa come stima del valore di aspettazione della popolazione.

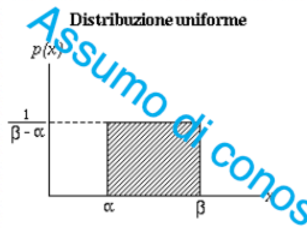
$x=\mu+\delta$

*Posso considerare x come la variabile aleatoria associata al processo di misura dipendente da un'altra variabile aleatoria ($\delta$) attraverso una relazione lineare.*
$E(x)=E(\mu+\delta)=\mu+E(\delta)=\mu$ se la popolazione errori di misura ha valore di aspettazione nullo.



Spazio Campione degli errori di misura

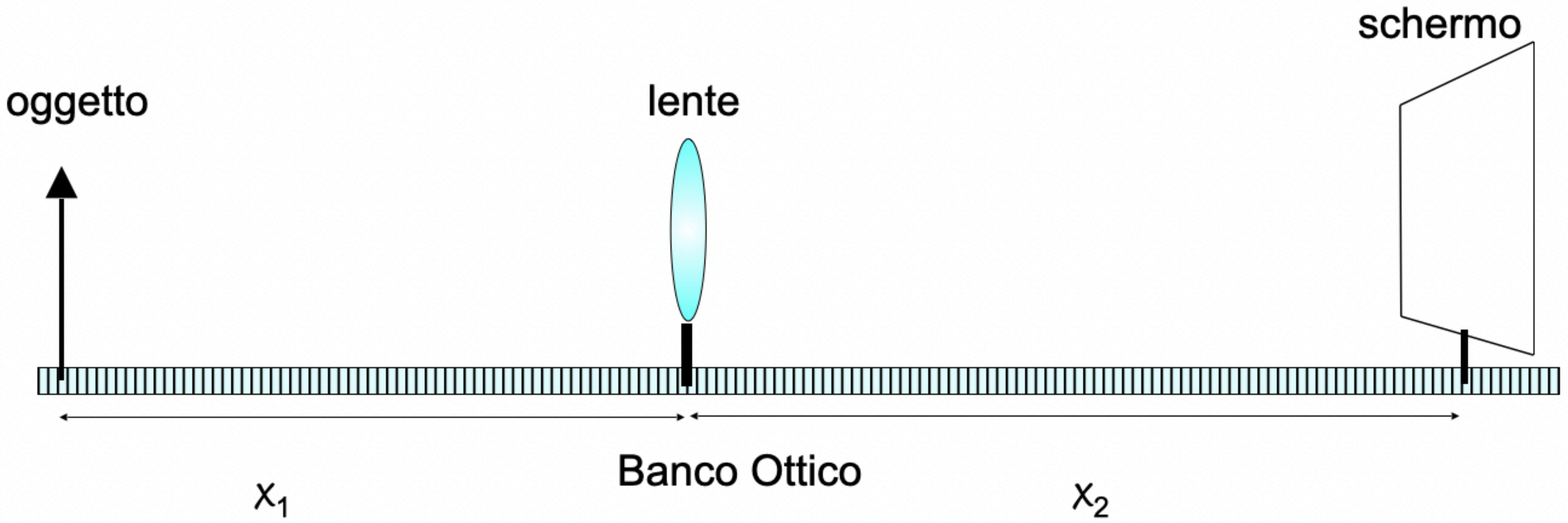
*18.2 +/- 0.3 cm*




Assumo di conoscerlo

# TEORIA DEGLI ERRORI DI MISURA

Esempio: Misura della lunghezza focale di una lente.



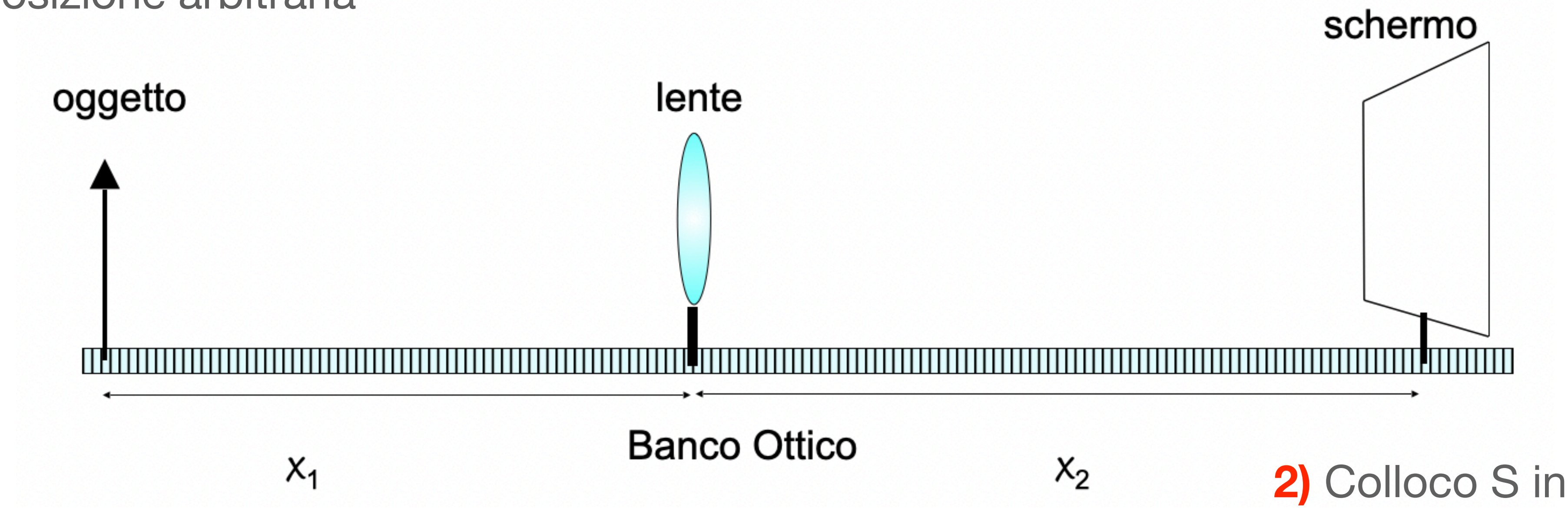

Per determinare la lunghezza focale della lente occorre misurare simultaneamente la distanza dalla lente dell'oggetto ($x_1$) e la distanza dalla lente dell'immagine ($x_2$).

$$\frac{1}{f} = \frac{1}{x_1} + \frac{1}{x_2}$$

$f$ lunghezza focale

## Procedura

**1)** Colloco O in posizione arbitraria




**2)** Colloco S in posizione arbitraria

**3)** Colloco L al centro dell'intervallino (ampio 2 cm) che consente di focalizzare I sullo schermo

Errori su $x_1$ e $x_2$:

Determino la posizione dell'oggetto $P_o$, la posizione della lente $P_l$ e la posizione dello schermo $P_s$.
Calcolo $x_1$ e $x_2$ per differenza tra le posizioni.
Ogni posizione ha un suo errore di misura. Calcolo con la propagazione l'errore su le distanza oggetto-lente e lente-schermo.

$$P_0 = 0.00 \pm 0.05 \quad cm$$

$$P_l = 30 \pm 1 \quad cm$$

$$P_s = 72.00 \pm 0.05 \quad cm$$

$$x_1 = 30.0 \pm 0.3$$

$$x_2 = 42.0 \pm 0.3$$

$$f = \frac{x_1 x_2}{x_1 + x_2} = 17.5$$

$$\sigma_f^2 = \left(\frac{\partial f}{\partial x_1}\right)^2 \sigma^2_{x_1} + \left(\frac{\partial f}{\partial x_2}\right)^2 \sigma^2_{x_2} \qquad \boxed{???}$$

# TEORIA DEGLI ERRORI DI MISURA

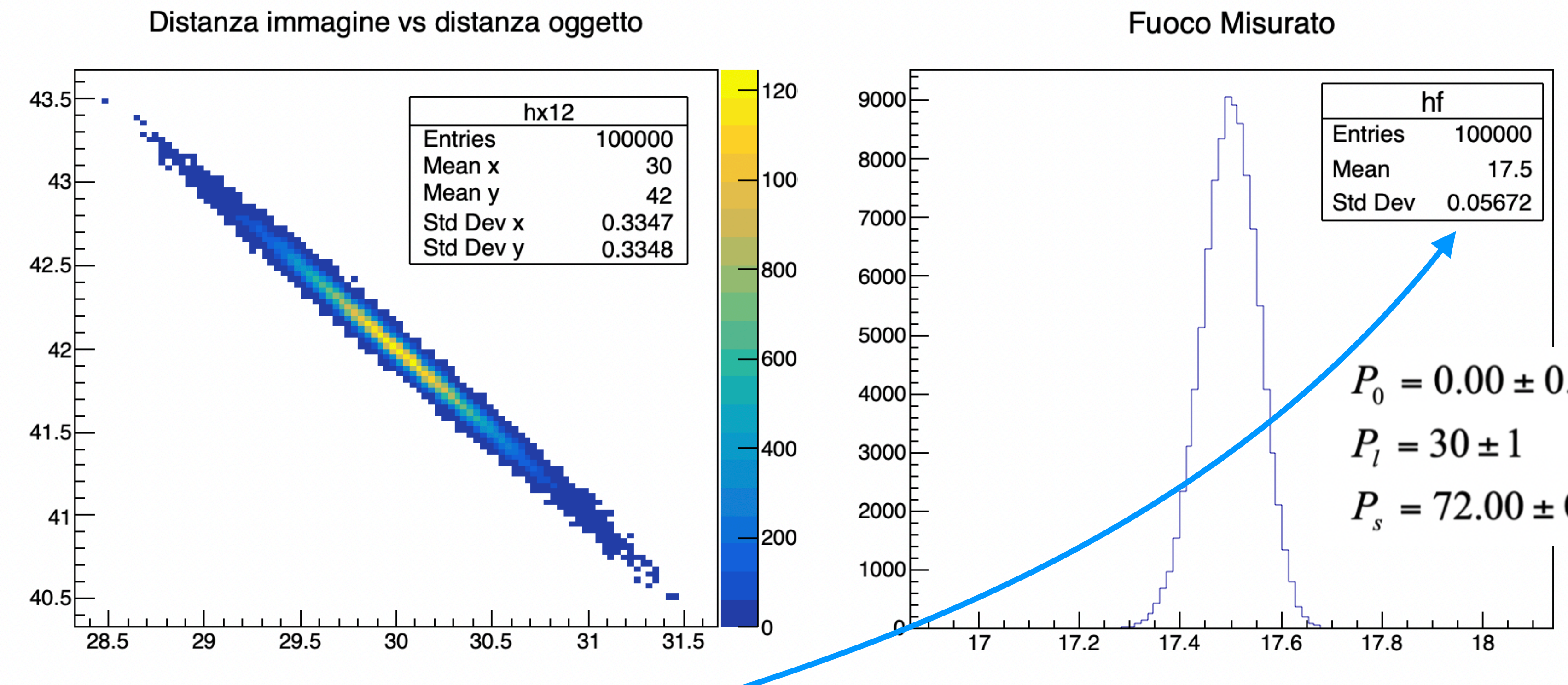
Errori massimi ed errori statistici

$$\sigma_{P_o} = \frac{0.05}{\sqrt{3}} = 0.03 \qquad \sigma_{P_s} = \frac{0.05}{\sqrt{3}} = 0.03 \qquad \sigma_{P_l} = \frac{1}{3}$$

$$\sigma_{x_1} = 0.3 \qquad \sigma_{x_2} = 0.3$$

$$\sigma_f^2 = \left(\frac{x_2(x_1 + x_2) - x_1 x_2}{(x_1 + x_2)^2}\right)^2 \sigma_{x_1}^2 + \left(\frac{x_1(x_1 + x_2) - x_1 x_2}{(x_1 + x_2)^2}\right)^2 \sigma_{x_2}^2$$

$$\sigma_f \cong 0.13$$




$P_0 = 0.00 \pm 0.05 \quad cm$

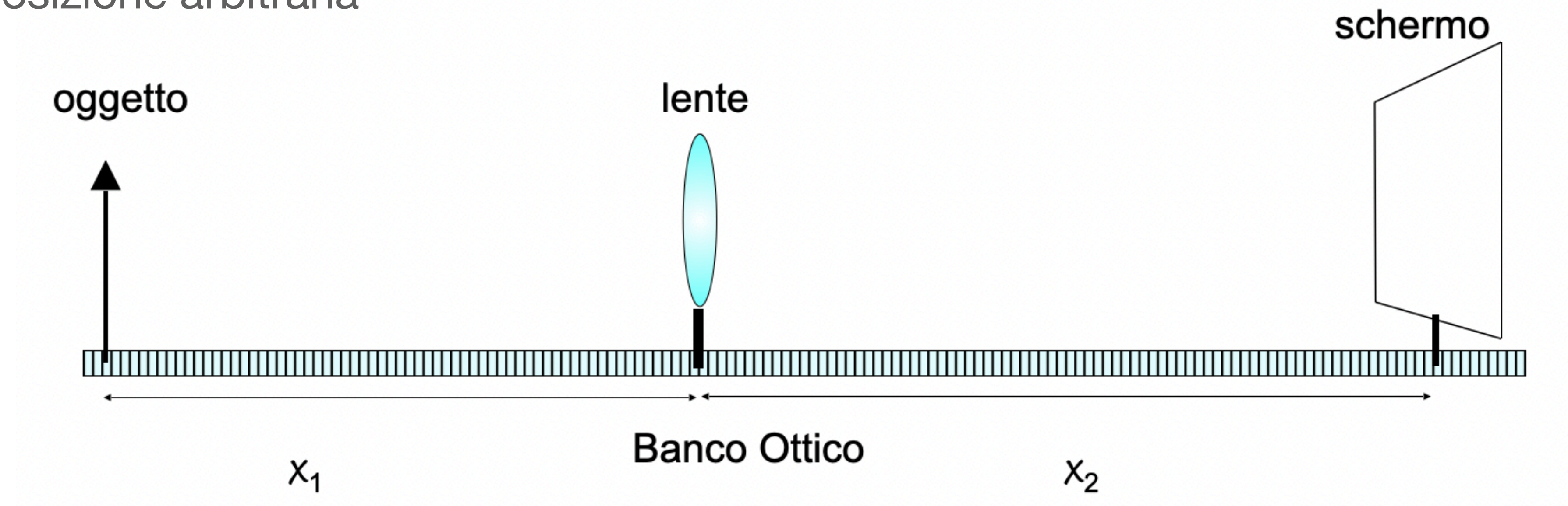
$P_l = 30 \pm 1 \quad cm$

$P_s = 72.00 \pm 0.05 \quad cm$

**Calcolo Analitico**
x1=30.000000 +/- 0.334581
x2=42.000000 +/- 0.334581
fuoco=17.500000
**sigma=0.127812** infatti la **var(f)=**Df_x1*var(x1)+Df_x2*var(x2))=**0.016336** <<<<<———— ignorando la correlazione
**Verifica Simulazione**
media di X1=29.998597, sigma(X1)=0.334730 e Var(x1)=0.112044
media di X2=42.001437, sigma(X2)=0.334760 e Var(x2)=0.112064
covarianza -0.111218
**Calcolo della varianza e sigma di F tenendo conto delle correlazione**
sigma=0.056528 var(f)=Df_x1*var(x1)+Df_x2*var(x2)+2Df_x1*Df_x2*cov(X1,X2)=**0.003195**
Variabile aleatoria derivata. Risultati Simulazione
media di f=17.498222, **sigma(f)=0.056716** e Var(f)=**0.003217**

Una diversa procedura

1) Colloco O in posizione arbitraria




2) Colloco L in posizione arbitraria

3) Colloco S al centro dell'intervallino (ampio 2 cm) che consente di focalizzare I sullo schermo

NOTA: ora $\sigma(x_2) \gg \sigma(x_1) \gg \sigma(P_L)$

Errori massimi ed errori statistici

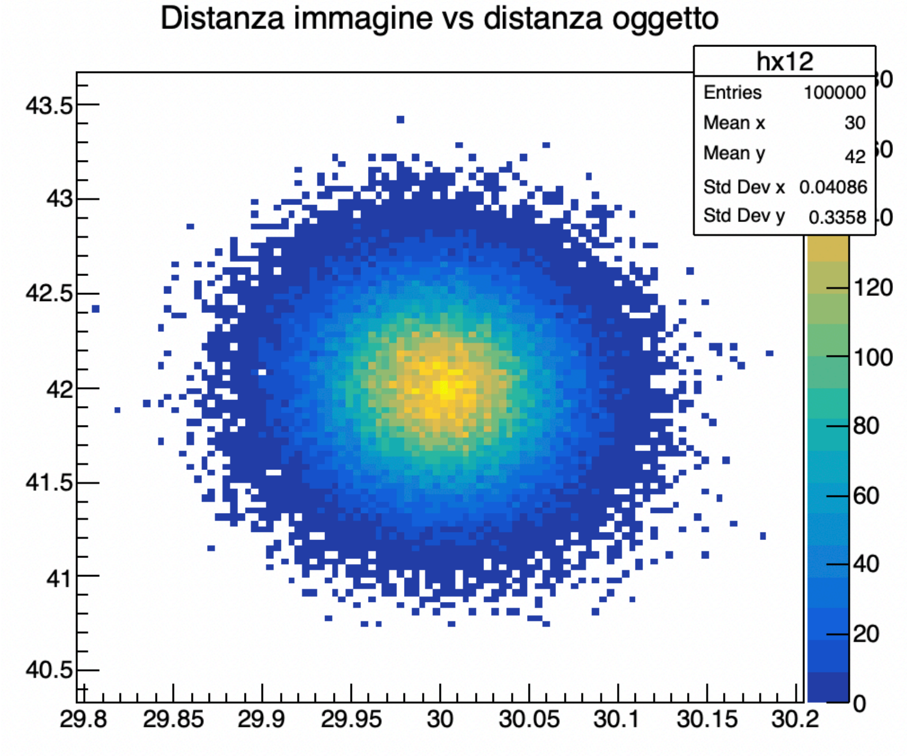
σ(PL ) σ(PS )

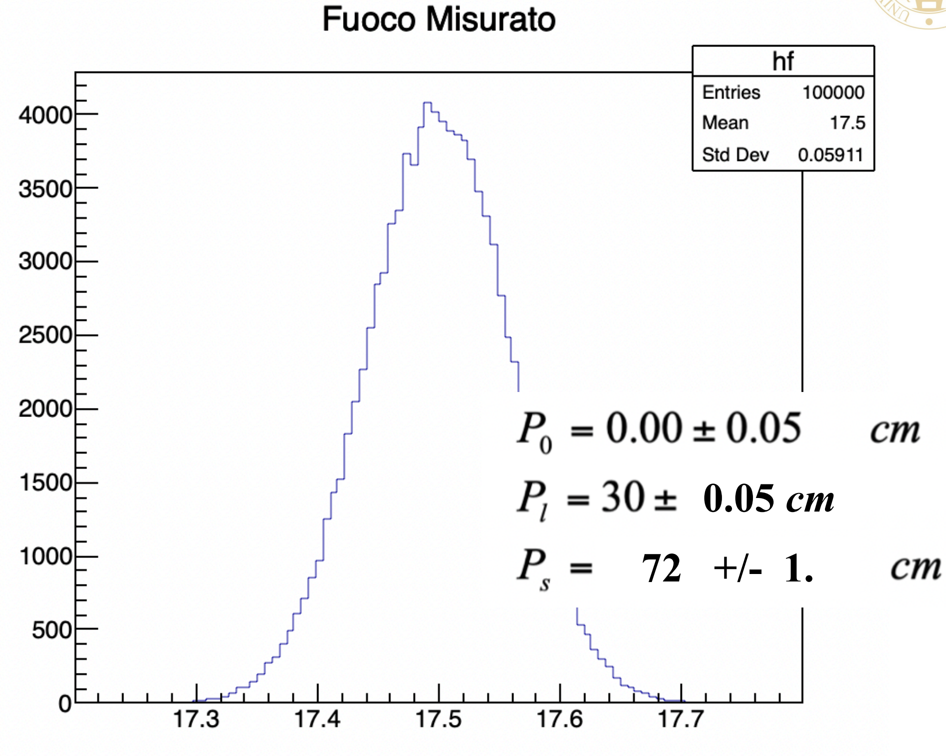
$$\sigma_{P_o} = \frac{0.05}{\sqrt{3}} = 0.03 \qquad \sigma_{P_s} = \frac{0.05}{\sqrt{3}} = 0.03 \qquad \sigma_{P_l} = \frac{1}{3}$$

$$\sigma_{x_1} = 0.045 \qquad \sigma_{x_2} = 0.3$$

$$\sigma_f^2 = \left(\frac{x_2(x_1+x_2) - x_1x_2}{(x_1+x_2)^2}\right)^2 \sigma_{x_1}^2 + \left(\frac{x_1(x_1+x_2) - x_1x_2}{(x_1+x_2)^2}\right)^2 \sigma_{x_2}^2$$

$$\sigma_f \cong 0.06$$




**Calcolo Analitico**
x1=30.000000 +/- 0.040825
x2=42.000000 +/- 0.334581
fuoco=17.500000
sigma=**0.059725** var(f)=Df_x1*var(x1)+Df_x2*var(x2))=***0.003567***
**Verifica Simulazione**
media di X1=29.999951, sigma(X1)=0.040861 e Var(x1)=0.001670
media di X2=41.999606, sigma(X2)=0.335764 e Var(x2)=0.112738
*covarianza -0.000833* *<<<<−−−−−−−−−−−−−−−−−−−− piccola covarianza*
**Calcolo della varianza e sigma di F tenendo conto delle correlazione**
sigma=0.058896 var(f)=Df_x1*var(x1)+Df_x2*var(x2)+2Df_x1*Df_x2*cov(X1,X2)=***0.003469***
Variabile aleatoria derivata. Risultati Simulazione
media di f=17.499630, sigma(f)=**0.059107** e Var(f)=***0.003494***

# ROOT EXTRA

- Object ownership https://root.cern.ch/root/htmldoc/guides/users-guide/ObjectOwnership.html

An object has ownership of another object if it has permission to delete it.

- ## 10.1 Ownership by Current Directory (gDirectory)

When a histogram, tree, or event list ( `TEventList` ) is created, it is added to the list of objects in the current directory by default. You can get the list of objects in a directory and retrieve a pointer to a specific object with the `GetList` method. This example retrieves a histogram.

```
TH1F *h = (TH1F*)gDirectory->GetList()->FindObject("myHist");
```

The method `TDirectory::GetList()` returns a `TList` of objects in the directory. It looks in memory, and is implemented in all ROOT collections. You can change the directory of a histogram, tree, or event list with the `SetDirectory` method. Here we use a histogram for an example, but the same applies to trees and event lists.

```
h->SetDirectory(newDir);
```

You can also remove a histogram from a directory by using `SetDirectory(0)` . Once a histogram is removed from the directory, it will not be deleted when the directory is closed. It is now your responsibility to delete this histogram once you have finished with it. To change the default that automatically adds the histogram to the current directory, you can call the static function:

```
TH1::AddDirectory(kFALSE);
```

Not all histograms created here after will be added to the current directory. In this case, you own all histogram objects and you will need to delete them and clean up the references. You can still set the directory of a histogram by calling `SetDirectory` once it has been created as described above.

Note that, when a file goes out of scope or is closed all objects on its object list are deleted.

# MACRO2.CPP

- The empty canvas problem

```
fumetto:23:33 [ ~/Documents/tmpMSC2024 ] root -l macro2.cpp
root [0]
Processing macro2.cpp...
Error in <TApplication::TApplication>: only one instance of TApplication allowed
Info in <macro2>: File opened successfully.
Info in <macro2>: Histogram h_gruppoI found.
Info in <TCanvas::MakeDefCanvas>:  created default TCanvas with name c1
```

```
// Apri il file ROOT e leggi l'istogramma del gruppo I
TFile *f = TFile::Open("InputPdf.root");  //<<<<-- Current directory: InputPdf.root:/
TH1D* h_GRB = (TH1D*)f->Get("h_gruppoI"); //<<<<-- h_GRB appartiene alla dir InputPdf.root:/ => e' distrutto
                                          //<<<<-- alla chiusura del file


// costruisce e riempie l'isto della pdf cumulativa
cumulativeHist->Draw();    //<<<<-- cumulativeHist appartiene alla current dir InputPdf.root:/ => e' distrutto
                          //<<<<-- alla chiusura del file

// riempie un istogramma Hist con dati selezionati con il metodo dell'inversione
// alla fine
Hist->Draw(); //<<<<-- Hist appartiene alla current dir InputPdf.root:/ => e' distrutto
              //<<<<-- alla chiusura del file

f->Close();   //<<<<-- file chiesto => istogrammi distrutti -> Canvas vuota
```

# MACRO2.CPP CURA n I

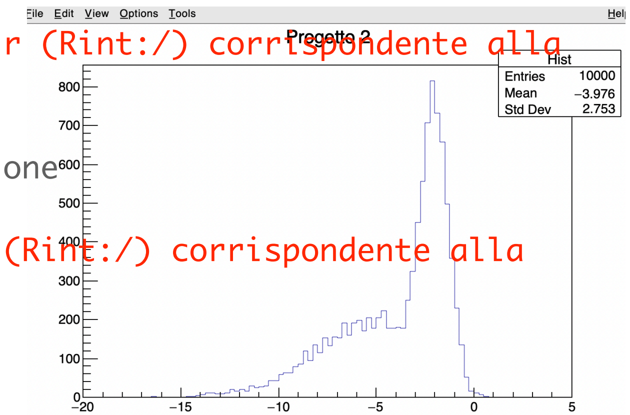
- The empty canvas problem

```
fumetto:23:33 [ ~/Documents/tmpMSC2024 ] root -l macro2.cpp
root [0]
Processing macro2.cpp...
Error in <TApplication::TApplication>: only one instance of TApplication allowed
Info in <macro2>: File opened successfully.
Info in <macro2>: Histogram h_gruppoI found.
Info in <TCanvas::MakeDefCanvas>:  created default TCanvas with name c1
```

```
// Apri il file ROOT e leggi l'istogramma del gruppo I
TFile *f = TFile::Open("InputPdf.root");  //<<<<-- Current directory: InputPdf.root:/
TH1D* h_GRB = (TH1D*)f->Get("h_gruppoI"); //<<<<-- h_GRB appartiene alla dir InputPdf.root:/ => e' distrutto
                                          //<<<<-- alla chiusura del file
h_GRB->SetDirectory(0);    // sottraiamo h_GRB alla "proprieta' della directory InputPdf.root:/" => non sara'
distrutto alla chiusura del file
f->Close(); // chiudiamo il file; h_GRB continua ad essere disponibile


// costruisce e riempie l'isto della pdf cumulativa
cumulativeHist->Draw();    //<<<<-- cumulativeHist appartiene alla current dir (Rint:/) corrispondente alla
sessione interattiva di root; sara' chiusa solo chiudendo root

// riempie un istogramma Hist con dati selezionati con il metodo dell'inversione
// alla fine
Hist->Draw();             //<<<<-- cumulativeHist appartiene alla current dir (Rint:/) corrispondente alla
sessione interattiva di root; sara' chiusa solo chiudendo root
```



# MACRO2.CPP CURA n I: **ATTENZIONE**

- The empty canvas problem

```
fumetto:23:33 [ ~/Documents/tmpMSC2024 ] root -l macro2.cpp
root [0]
Processing macro2.cpp...
Error in <TApplication::TApplication>: only one instance of TApplication allowed
Info in <macro2>: File opened successfully.
Info in <macro2>: Histogram h_gruppoI found.
Info in <TCanvas::MakeDefCanvas>:  created default TCanvas with name c1
```

```
// Apri il file ROOT e leggi l'istogramma del gruppo I
TFile *f = TFile::Open("InputPdf.root");  //<<<<-- Current directory: InputPdf.root:/
TH1D* h_GRB = (TH1D*)f->Get("h_gruppoI"); //<<<<-- h_GRB appartiene alla dir InputPdf.root:/ => e' distrutto
                                          //<<<<-- alla chiusura del file
h_GRB->SetDirectory(0);    // sottraiamo h_GRB alla "proprieta' della directory InputPdf.root:/" => non sara'
distrutto alla chiusura del file
///////f->Close(); // NON chiudiamo il file



// costruisce e riempie l'isto della pdf cumulativa
cumulativeHis-t>Draw();    //<<<<-- cumulativeHist appartiene alla current dir InputPdf.root:/ => e' distrutto
                                          //<<<<-- alla chiusura del file

// riempie un istogramma Hist con dati selezionati con il metodo dell'inversione
// alla fine
Hist->Draw();                //<<<<-- cumulativeHist appartiene alla current dir InputPdf.root:/ => e' distrutto
                                          //<<<<-- alla chiusura del file

f->Close(); // chiudiamo il file; h_GRB continua ad essere disponibile; gli altri istogrammi NO
```

## CANVAS vuota

# MACRO2.CPP CURA n 2

- The empty canvas problem

```
fumetto:23:33 [ ~/Documents/tmpMSC2024 ] root -l macro2.cpp
root [0]
Processing macro2.cpp...
Error in <TApplication::TApplication>: only one instance of TApplication allowed
Info in <macro2>: File opened successfully.
Info in <macro2>: Histogram h_gruppoI found.
Info in <TCanvas::MakeDefCanvas>:  created default TCanvas with name c1
```

```
TH1::AddDirectory(kFALSE);
// Apri il file ROOT e leggi l'istogramma del gruppo I
TFile *f = TFile::Open("InputPdf.root");  //<<<<-- Current directory: InputPdf.root:/
TH1D* h_GRB = (TH1D*)f->Get("h_gruppoI"); //<<<<-- h_GRB NON appartiene alla dir InputPdf.root:/ => e' distrutt
                                          //<<<<-- alla chiusura del file


// costruisce e riempie l'isto della pdf cumulativa
cumulativeHist->Draw();    //<<<<-- cumulativeHist appartiene alla current dir InputPdf.root:/ => e' distrutto
                          //<<<<-- alla chiusura del file

// riempie un istogramma Hist con dati selezionati con il metodo dell'inversione
// alla fine
Hist->Draw(); //<<<<-- Hist appartiene alla current dir InputPdf.root:/ => e' distrutto
              //<<<<-- alla chiusura del file

f->Close();   //<<<<-- file chiesto => istogrammi distrutti -> Canvas vuota
```