Scuola Universitaria Professionale
della Svizzera Italiana

SUPSI

Dipartimento
Ambiente
Costruzioni e
Design

**Istituto
Scienze
della Terra**

Corso C111.01 - Geomatica

Teoria degli errori

Massimiliano Cannata

Rappresentazione di una misura di precisione

$$(x \pm \sigma_x)_u$$

x = misura

σ_x = incertezza della misura

u = unità di misura

Il problema degli errori in topografia

La misura (x) viene effettuata con uno strumento (S) che osserva una grandezza (G)

$$G \rightarrow S \rightarrow x$$

- La topografia è basata sulla misura di grandezze (angoli, distanze, etc.)
- Ogni misura è affetta da errori
- L'entità degli errori definisce l'incertezza della misure

Gestione dell’incertezza

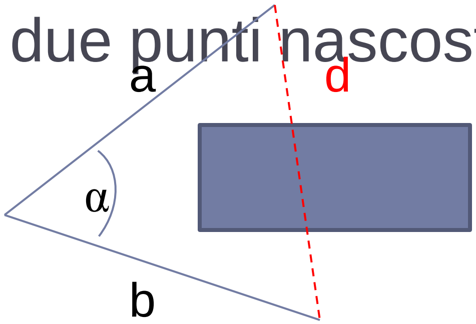
- Progettazione:
  - scelta dello strumento e del metodo di rilievo che mi garantisca la precisione richiesta
- In campagna:
  - riduzione delle fonti di errore per ottenere la massima precisione

Misure dirette e indirette

- Misure dirette:
  - La grandezza viene osservata direttamente tramite uno strumento (angolo α, distanze a e b)

- Misure indirette:
  - La grandezza viene derivata in funzione di altre grandezze osservate (distanza d tra due punti nascosti calcolata con teorema di Carnot)

$$d = \sqrt{a^2 + b^2 - 2ab\cos\alpha}$$



Misure dirette

$$G \rightarrow S \rightarrow x$$

- Il legame tra la misura e la grandezza deve essere univoco
- Lo strumento esegue misure discrete della grandezza e le discontinuità di misura sono dovute alla sensibilità dello strumento (più piccola unità misurabile)
- L'univocità è possibile solo all'interno della massima grandezza misurabile dallo strumento
- Bontà di uno strumento è la capacità di generare misure univoche

Variabilità delle misure ripetute

$$G \rightarrow S \rightarrow x$$

- Sperimentalmente si nota che misure ripetute sulla stessa grandezza generano valori diversi tra loro

$$G \rightarrow S \rightarrow X$$

- G=grandezza,
- S=strumento,
- X=popolazione delle misure possibili

Cause della variabilità delle misure

- Instabilità dello strumento
- Incapacità di definire la grandezza osservata
- Variabilità della grandezza nel tempo
- Uso inappropriato dello strumento

Instabilità strumentale

- La stabilità di uno strumento di misura è la sua attitudine a mantenere costanti le sue caratteristiche

- Uno strumento ideale, una volta realizzato, dovrebbe mantenere le sue caratteristiche costanti.

- Nella realtà, per cause interne (es. deperimento e consumo dei componenti, difettosità nella realizzazione) o esterne (es. variazione delle condizioni ambientali, cattivo uso), la strumentazione varia le proprie caratteristiche in maniera più o meno significativa.

Instabilità per cause esterne

- Imputabile all'interazione strumento-ambiente:
  - in particolare alla variabilità dei parametri ambientali (temperatura, umidità, vento, etc.: u, v, w,...t) dell'unità di misura adottata (U) e della quantità di grandezza da misurare (G)

$$X = f\,(G/U, u, v, w, \ldots t)$$

- La variabilità dei parametri ambientali può essere assunta come una perturbazione casuale, e di conseguenza anche le misure ripetute possono essere considerate come estrazioni casuali tra quelle possibili.

Misura vera

- La misura vera è quella misura è quello che vogliamo misurare (il valore univoco di quella grandezza)

- Le perturbazioni, nelle ipotesi fatte sulla loro casualità, oscilleranno nel limitato periodo relativo alla ripetizione delle misure nell'intorno di un livello medio, ripassando più frequentemente su questo valore, che risulta di conseguenza anche il più probabile.

- La definizione della misura vera (X_0) non può essere che legata al valore più probabile:

$$X_0 = f\,(G/U,\ u_m,\ v_m,\ w_m, \ldots t_m)$$

Errore di misura

- La definizione di una misura vera all'interno della popolazione delle misure possibili lega il suo valore alla sensibilità dello strumento.
- La misura vera, anche se definita, resta ugualmente sconosciuta e si parla più propriamente di misura vera convenzionale.
- L'aspetto pratico più utile della sua esistenza è quello di rappresentare il riferimento per la definizione dell'errore di misura:

$$\varepsilon = X - X_0$$

Tipologie di errori di misura

- Poichè le cause perturbatrici sono sempre presenti le misure sono sempre affette da errori.

- Si distinguono errori:
  - grossolani,
  - sistematici
  - accidentali

$$\varepsilon = \varepsilon_g + \varepsilon_s + \varepsilon_a$$

Gli errori grossolani (ε_g)

- Sono quegli errori dovuti principalmente alla distrazione dell'operatore (non messa in bolla dello strumento, errore di trascrizione, errori di codifica dei punti, etc.)

- Questi errori devono essere eliminati prima dell'elaborazione delle misure tramite strumenti di controllo
  - ad esempio per gli angoli si verifica che la somma degli angoli interni ad un triangolo sia pari a 200 gon

Gli errori sistematici (ε_s)

- Sono quegli errori dovunti ad un difetto dello strumento o ad effetto fisico che causano delle misure sfalsate:
  - di una costante: ad esempio se lo zero strumentale non è settato correttamente
  - di una deriva: ad esempio se l'errore è in funzione della distanza 1ppm

- Se si conoscono le cause si possono eliminare applicando delle correzioni opportune

Gli errori accidentali (ε_a)

- Sono la parte casuale dell'errore dovuta alla somma di cause molteplici non note
- Tali errori hanno una distribuzione casuale
- Non si possono eliminare, ma se ne può limitare gli effetti
- D'ora in avanti si supporrà sempre che:
  - Gli errori grossolani siano stati eliminati
  - Le cause dei possibili errori sistematici siano note e quindi sia possibile ridurne gli effetti

Teorema centrale della statistica

- Se alla formazione di un fenomeno, come quello della variabilità degli errori accidentali, concorrono molte cause, di cui nessuna preponderante, a cui sono associate variabili casuali a varianza non nulla, la distribuzione probabilistica del fenomeno tende ad una distribuzione normale.

Parametri statistici

- Si introducono alcuni parametri statistici per ottenere informazioni sintetiche sulla popolazione

  - Indice di posizione: media

$$\mu_X = \sum_i^m X_i \frac{F_i}{N}$$

  - Indice di dispersione: s.q.m. (e varianza)

$$\sigma_X = \pm \sqrt{\sum_i^m \left(X_i - \mu_X\right)^2 \frac{F_i}{N}}$$

Variabile casuale discreta

- È descritta da una serie di valori con un associata probabilità

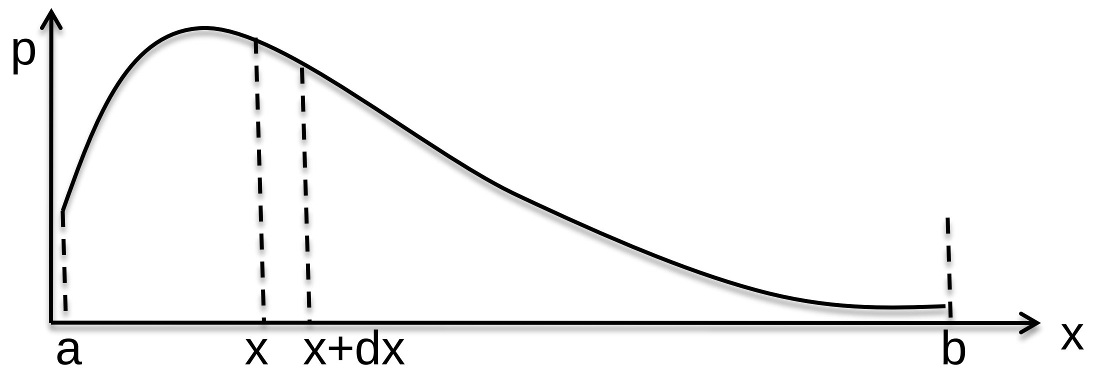
$$X\begin{cases} X_1, X_2, \ldots, X_i, \ldots, X_m \\ p_1,\ p_2, \ldots, p_i, \ldots, p_m \end{cases} \quad con \sum p_i = 1$$

- Il comportamento della variabile è descritto dai suoi momenti fondamentali: media e varianza (o scarto quadratico medio s.q.m. pari alla radice della varia

$$\mu_X = \sum_i^m X_i p_i \qquad\qquad \sigma_X = \pm\sqrt{\sum_i^m \left(X_i - \mu_X\right)^2 p_i}$$

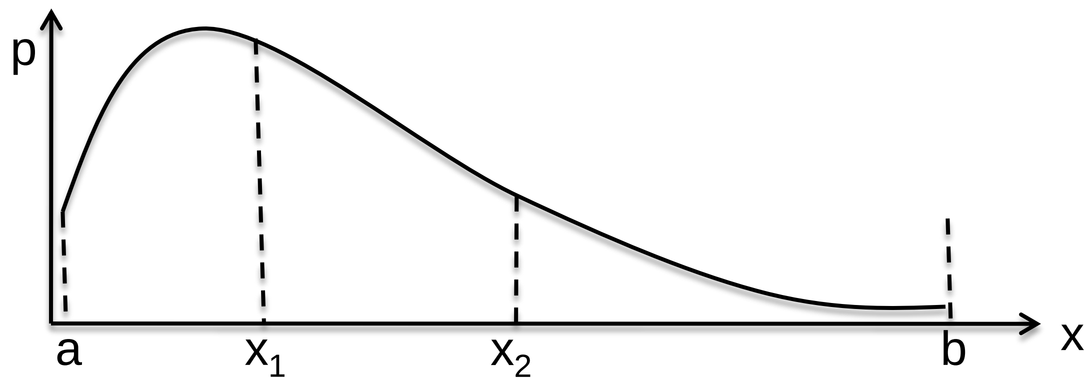
Variabile casuale continua

- È descritta da una funzione continua di **funzione di densità di probabilità** (o **pdf** dall'inglese probability density function)



- La probabilità elementare di un evento compreso tra *x* e *x+dx* è rappresentato dall'area sottesa da questo intervallo infinitesimo: $dp = f(x)\, dx$

Variabile casuale continua



- La probabilità elementare di un evento compreso tra x_1 e x_2 è rappresentato integrale:

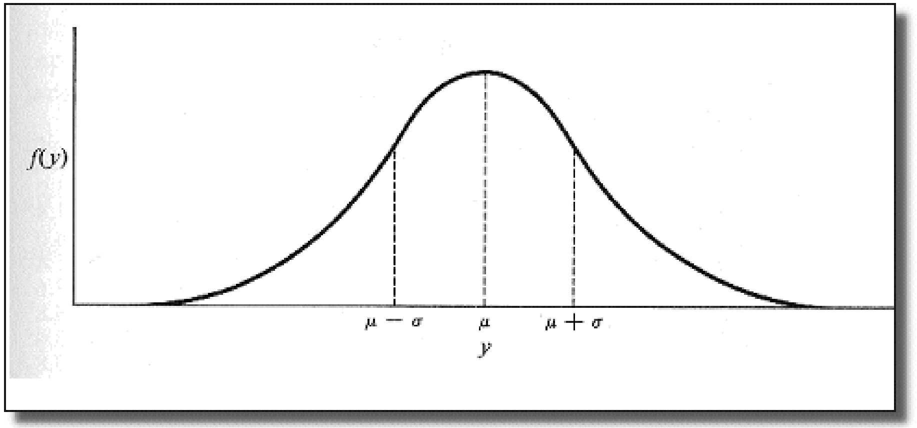
$$p = \int_{x_1}^{x_2} f(x)dx$$

- Mentre la probabilità totale risulta pari a 1:

$$p = \int_{-\infty}^{+\infty} f(x)dx = 1$$

Curva di Gauss

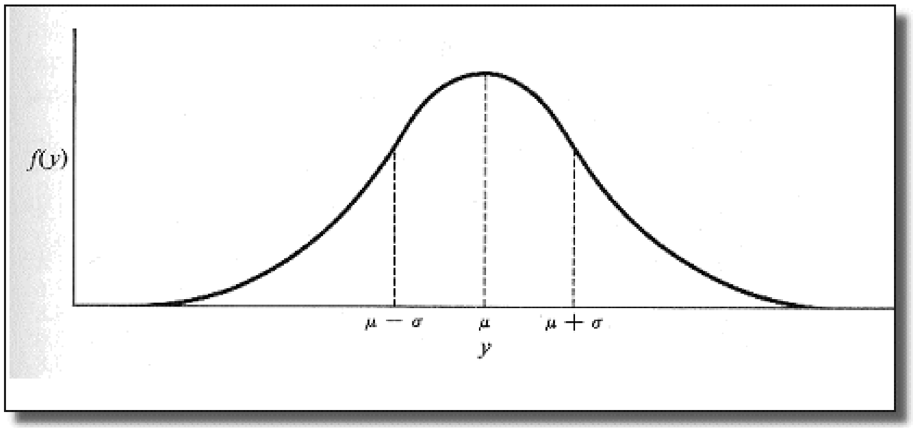
- Le variabili casuali normali (nel nostro caso delle misure) sono descritte da una funzione di densità di probabilità detta di Gauss:



- Descritta dalla funzione:

$$p_X(x) = \frac{1}{\sqrt{2\pi\sigma^2}}\, e^{-\frac{(x-\mu)^2}{2\sigma^2}} \quad \text{con} \quad -\infty < x < \infty$$

Curva di Gauss



- La curva è caratterizzata da media e varianza descritte dalle seguenti espressioni:

$$\mu_X = \int_{-\infty}^{+\infty} x f(x) dx \qquad \sigma_X^2 = \int_{-\infty}^{+\infty} (x - \mu_X)^2 f(x) dx$$

Media empirica semplice

- Non <u>si può conoscere il valore medio </u>poichè a priori non si conosce la popolazione delle misure possibili, quindi il valore medio della gaussiana deve essere stimato attraverso la media empirica calcolata sulle misure ripetute effettuate.

- Questo può essere fatto poichè la legge empirica del caso ci garantisce che all'aumentare delle misure ripetute (nel caso di errori accidentali e quindi casuali) <u>la media empirica tende al valore della media teorica</u>

Media empirica semplice

- Data una serie di misure ripetute (x_1, x_2, x_3, … , x_n) la media empirica semplice (come stima della media teorica) è data da:

$$M = \frac{1}{n}\sum_{i=1}^{n} x_i$$

- Questa media <u>M va pensata come un valore casuale</u> estratto da una serie di misure casuali (misure ripetute) così come x_i è un valore estratto a caso tra tutte le misure possibili X

Incertezza della media empirica

- Il valore teorico σ_X è incognito, ma può essere stimato attraverso il valore empirico σ'_X dato dall'espressione:

$$\sigma'_X = \pm\sqrt{\frac{\sum_{i=1}^{n}(x_i - M)^2}{n-1}}$$

- In definitiva l'incertezza della media empirica semplice è valutata attraverso il valore empirico dato all'espressione:

$$\sigma'_M = \frac{\sigma'_X}{\sqrt{n}}$$

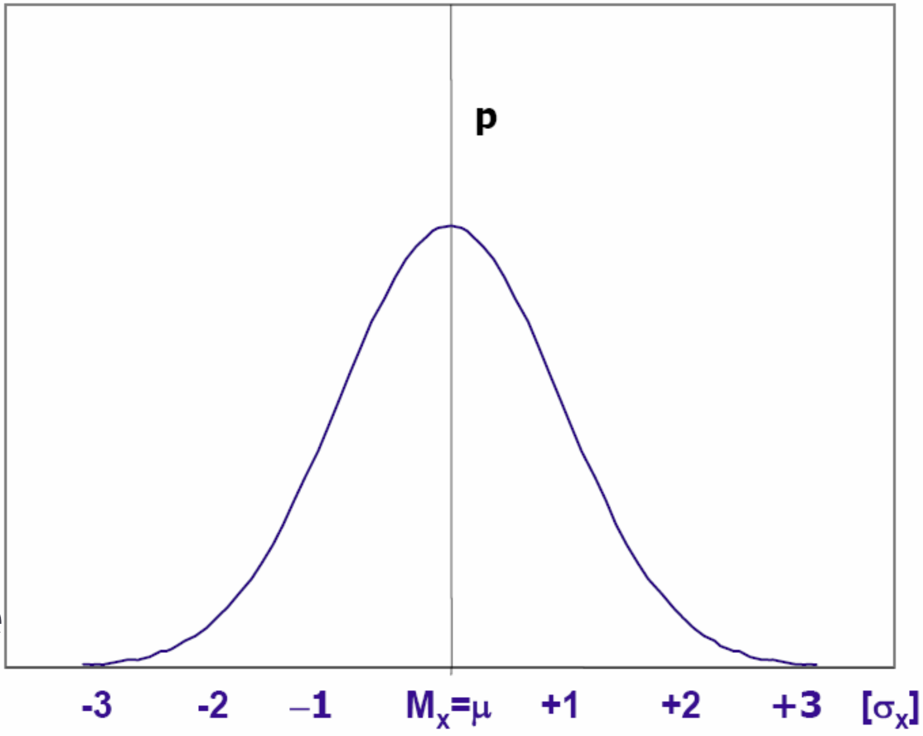
- Il risultato della misure sarà quindi in definitiva:

$$(M \pm \sigma'_M)_u$$

Considerazioni sul risultato di una misura

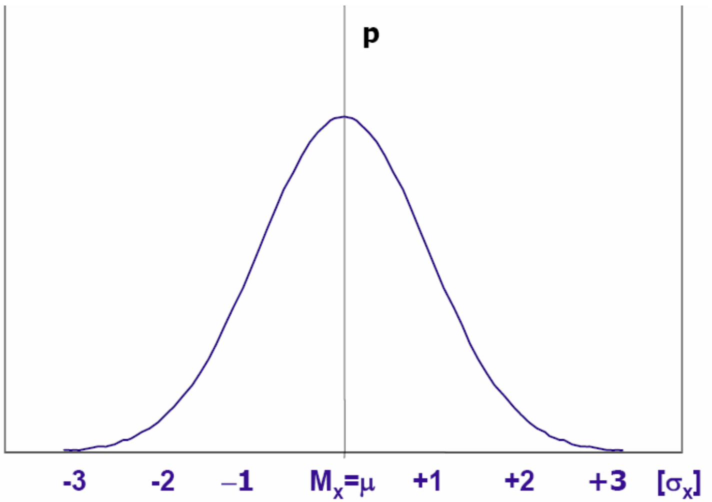
- La popolazione di misure possibili X ha una distribuzione normale ed è compresa in percentuali diverse e con diverse probabilità in diversi intervalli dalla media.

- Ad esempio:
  - il 68% delle misure è compreso tra M $\pm\sigma_x$
  - il 95% delle misure è compreso tra M $\pm 2\sigma_x$
  - il 99,7% delle misure è compreso tra M $\pm 3\sigma_x$



Considerazioni sul risultato di una misura

- Analogamente la popolazione delle medie empiriche possibili Y è distribuita normalmente e si verifica dunque che:
  - il 68% delle misure è compreso tra $\mu \pm \sigma_M$
  - il 95% delle misure è compreso tra $\mu \pm 2\sigma_M$
  - il 99,7% delle misure è compreso tra $\mu \pm 3\sigma_M$



- Questo significa ad esempio che la media M ha una probabilità pari a 0,95 di trovarsi nell'intorno di $\pm 2\sigma_M$ dalla media teorica μ ($M_Y=\mu$), ovvero che **la misura vera ha una probabilità pari a 0,95 di cadere nell'intorno della media empirica semplice M**

Esempio numerico

- Si consideri il seguente campione costituito dalla misura di un angolo ripetuta 5 volte:
  - α = 300,450 ; 300,458 ; 300,462 ; 300,443 ; 300,440 gon
- Calcolo della media empirica M_α:
  - $M_\alpha = \Sigma\, \alpha_i / n = 1502{,}253 / 5 = 300{,}4506$ gon
- Calcolo dello s.q.m. del campione:
  - $\sigma'_X = \pm\sqrt{\frac{\sum_{i=1}^{n}(x_i - M)^2}{n-1}}$ $= 9{,}423 \cdot 10^{-3}$ gon
- Calcolo delle s.q.m. della media empirica:
  - $\sigma'_M = \frac{\sigma'_X}{\sqrt{n}}$ $= = 4{,}214 \cdot 10^{-3}$ gon
- Stima del valore vero dell'angolo:
  - $\alpha = M_\alpha \pm \sigma'_M = (300{,}4506 \pm 0{,}0042)$ gon

Esercizio

- Stimare il valore vero di una distanza sulla base delle seguenti misure ripetute:

  - D = 102,56 ; 102,49 ; 102,98 ; 103,02 ; 101,89 ; 102,34

S.q.m. e tolleranza

- Nei capitolati d'appalto per la cartografia, nei quali è prescritta l'incertezza con cui devono essere eseguite le misure compare il termine **tolleranza**
- Nella schede tecniche di uno strumento topografico viene invece utilizzato lo **s.q.m.** (σ) per indicare l'incertezza delle misure che è in grado di produrre

Normalmente per tolleranza si intende un intervallo nell'intorno della media di ±3σ

- NB: se è richiesta una tolleranza pari a t, si dovrà operare con uno strumento che garantisca un'incertezza pari a 1/3 di t

Media empirica ponderata

- Se una stessa grandezza è <u>misurata con diversi gradi di precisione</u> (strumenti diversi, situazioni ambientali diverse, etc..) le misure vengono trattate utilizzando la media empirica ponderata:

$$M_p = \frac{\sum_{i=1}^{n} x_i \, p_i}{\sum_{i=1}^{n} p_i} \quad \text{essendo} \quad p_i = \frac{{\sigma_0}^2}{\sigma'_{x_i}}$$

- La stima della varianza dell'unità di peso (σ_0^2) è data da:

$${\sigma_0'}^2 = \frac{\sum_{i=1}^{n} (x_i - M_p)^2 \, p_i}{n-1}$$

Incertezza della media empirica ponderata

- L'incertezza della media ampirica ponderata (s.q.m.) viene calcolata con la seguente equazione:

$$\sigma_{M_p} = \sqrt{\frac{{\sigma_0'}^2}{\sum_{i=1}^{n} p_i}}$$

Misure indirette

- Una misura indiretta viene ottenuta attraverso la misura diretta o anche indiretta di altre grandezze X_i chiamate determinanti (ad esempio distanza ed angolo)
- Si chiama modello funzionale $Z = f(X_1, X_2, ..., dX_n)$ dove Z rappresenta la variabile casuale associata alla popolazione di misure possibili della misura indiretta il modello che porta alla determinazione della grandezza desiderata

Misura vera convenzionale

- Per il th. centrale della statistica, se le X_i sono distribuite normalmente, anche la misura di Z è distribuita normalmente (<u>se le X_i sono casuali lo è anche la Z</u>)

- Si definisce come misura vera convenzionale (Z_0) per la misura indiretta quella che si otterrebbe introducendo nel legame funzionale i valori veri delle grandezze determinanti $Z_0 = f(X_{0,1}, X_{0,2}, \ldots ,X_{0,n})$

Valor medio di una misura indiretta

- Restando sempre valida l'ipotesi di accidentalità degli errori delle grandezze determinanti, le misure vere delle determinanti coincidono con le medie delle variabili casuali legate alle rispettive popolazioni di misure possibili. Sulla base di questa ipotesi si può dimostrare che $Z_0 = M(z)$ che indichiamo con μ_z:

$$Z_0 = \mu_z = f(\mu_{X1}, \mu_{X2}, \ldots, \mu_{Xn})$$

- Nella pratica non si conoscono i valori teorici μ_{Xi}, ma solo le loro stime M_{Xi}. Si dimostra che introducendo questi valori nel legame funzionale della misura indiretta si ottiene per questa una stima accettabile:

$$M_z = f(M_{x1}, M_{x2}, \ldots, M_{xn})$$

S.q.m. di una misura indiretta

- Un valore empirico della stima dell'incertezza di una misura indiretta può essere ottenuto grazie alla <u>legge di propagazione degli errori</u>.
- Nell'ipotesi che <u>le grandezze determinanti</u> X_i <u>siano indipendenti</u> l'una dall'altra risulta:

$$\sigma'_Z = \pm\sqrt{\left(\frac{\partial f}{\partial X_1}\right)^2_{M_{x_{1,\dots,n}}} \cdot \sigma^2_{X_1} + \left(\frac{\partial f}{\partial X_2}\right)^2_{M_{x_{1,\dots,n}}} \cdot \sigma'^2_{X_2} + \dots + \left(\frac{\partial f}{\partial X_n}\right)^2_{M_{x_{1,\dots,n}}} \cdot \sigma'^2_{X_n}}$$

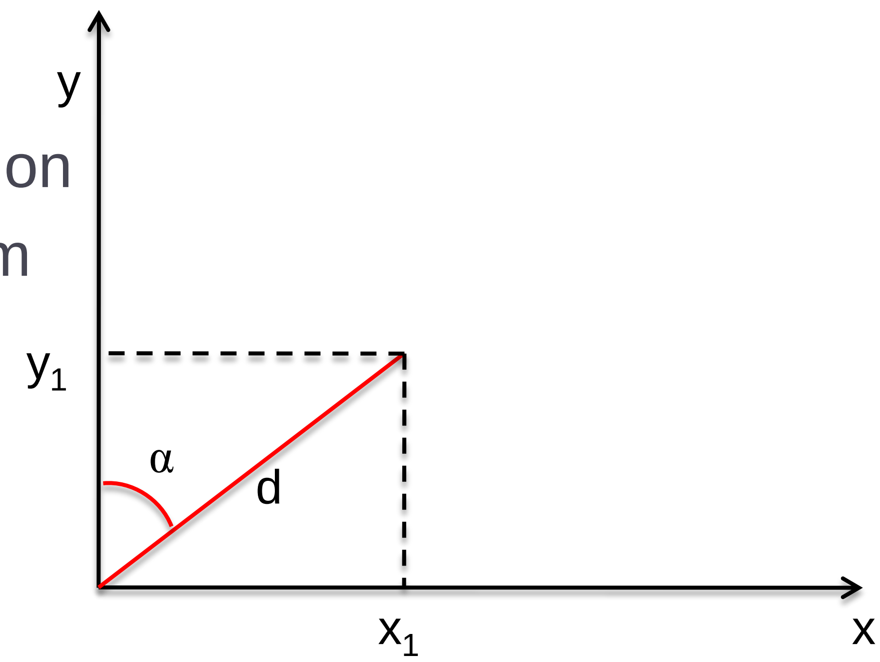
- L'entità dell'incertezza di z dipende dagli s.q.m. delle singole grandezze, ma anche dai valori numerici delle determinanti, in quanto le derivate sono calcolate sulla base di questi valori
- In caso di grandezze X_i dipendenti, occorre aggiungere alcuni termini correttivi (correlazioni)

Esempio numerico

- Calcolo dell'incertezza di un punto misurato per coordinate polari

- Misure:
  - α = 40,03 ; 40,07 ; 40, 04 gon
  - d = 50,27 ; 50,19 ; 50,32 ; m

- Relazioni funzionale:
  - $x_1 = d \sin \alpha$
  - $y_1 = d \cos \alpha$



Esempio numerico

- Calcolo della media empirica
  - $M_\alpha = \Sigma\, \alpha_i / n = 40{,}047$ gon
  - $M_d = \Sigma\, d_i / n = 50{,}26$ m
- Calcolo dello s.q.m. dei campioni: $\sigma'_X = \pm\sqrt{\frac{\sum_{i=1}^{n}(x_i - M)^2}{n-1}}$
  - $\sigma'_\alpha = 0.02082$ gon
  - $\sigma'_d = 0.0656$ m
- Calcolo delle s.q.m. della media empirica: $\sigma'_M = \frac{\sigma'_X}{\sqrt{n}}$
  - $\sigma'_{M\alpha} = 0.00787$ gon
  - $\sigma'_{Md} = 0.02478$ m
- Stima del valore vero dell’angolo:
  - $\alpha = 40{,}047 \pm 0.00787$ gon == 0.629 ± 0.000124 rad
  - $d = 50{,}26 \pm 0.02478$ m

Esempio numerico

- M_x = 50,26 sin (0.629) = 29.572
- M_y = 50,26 cos (0.629) = 40.639
- $\sigma_{x_1}^2 = \left(\frac{\partial x_1}{\partial \alpha}\right)^2 \sigma_\alpha^2 + \left(\frac{\partial x_1}{\partial d}\right)^2 \sigma_d^2 = (d\cos\alpha)^2 \sigma_\alpha^2 + (\sin\alpha)^2 \sigma_d^2$ = 0.000238
- $\sigma_{y_1}^2 = \left(\frac{\partial y_1}{\partial \alpha}\right)^2 \sigma_\alpha^2 + \left(\frac{\partial y_1}{\partial d}\right)^2 \sigma_d^2 = (-d\sin\alpha)^2 \sigma_\alpha^2 + (\cos\alpha)^2 \sigma_d^2$ = 0.000415
- σ'_{x1} = 0.0154
- σ'_{y1} = 0.0204
- Valori veri stimati: P (29.572 ± 0.0154, 40.639 ± 0.0204)