Note su applicazioni del Principio della Massima Verosimiglianza

Media aritmetica $\stackrel{?}{=}$ migliore stima del valore atteso

$x_1, x_2, \dots, x_N$ misure ripetute di x

supponiamo che si tratti di un campione estratto da una distribuzione limite di tipo "gaussiano"

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

Principio della massima verosimiglianza :

la migliore stima per μ sarà data da quel valore che rende massima la probabilità di osservare l'intero campione di N valori effettivamente trovati ripetendo le misure :

$$P(x_1, x_2, \dots, x_N) = P(x_1) \cdot P(x_2) \cdot \dots \cdot P(x_N)$$

$$\propto \frac{1}{\sigma} e^{-\frac{(x_1-\mu)^2}{2\sigma^2}} \cdot \frac{1}{\sigma} e^{-\frac{(x_2-\mu)^2}{2\sigma^2}} \cdot \dots \dots \frac{1}{\sigma} e^{-\frac{(x_N-\mu)^2}{2\sigma^2}}$$

$$= \frac{1}{\sigma^N} \exp\left[-\frac{\sum_1^N{}_i (x_i-\mu)^2}{2\sigma^2}\right]$$

$$\boxed{P(x_1, x_2, \dots, x_N) = max \longrightarrow \sum_1^N{}_i (x_i-\mu)^2 = \text{minimo al variare di } \mu}$$

$$0 = \frac{\partial}{\partial\mu}\left[\sum_1^N{}_i (x_i-\mu)^2\right] = \sum_1^N{}_i \left[2(x_i-\mu)(-1)\right] = -2 \sum_1^N{}_i (x_i-\mu) =$$

$$= -2\left(\sum_1^N{}_i x_i - N\mu\right) \longrightarrow \boxed{\mu = \sum_1^N{}_i x_i / N = \langle x \rangle}$$

Qual'è il miglior estimatore della deviazione standard ?

$$\sigma \overset{?}{=} \sqrt{\frac{\sum_1^N i(x_i-\mu)^2}{N}} \quad ; \quad \sigma \overset{?}{=} \sqrt{\frac{\sum_1^N i(x_i-\langle x\rangle)^2}{N-1}}$$

Se conosco μ, cioè non devo fare uso delle N misure per stimarlo tramite $\langle x\rangle = \sum_1^N i x_i / N$.

→ Dal principio della massima verosimiglianza,

$$\boxed{P(x_1, x_2, \ldots, x_N) \propto \frac{1}{\sigma^N} \exp\left[-\frac{\sum_1^N i(x_i-\mu)^2}{2\sigma^2}\right] = \text{massimo al variare di } \sigma}$$

$$0 = \frac{\partial}{\partial\sigma}\left\{\sigma^{-N} \cdot \exp\left(-\frac{\sum_1^N i(x_i-\mu)^2}{2\sigma^2}\right)\right\} =$$

$$= -N \cdot \sigma^{-N-1} \cdot \exp\left(-\frac{\sum_1^N i(x_i-\mu)^2}{2\sigma^2}\right) + \sigma^{-N} \cdot \left[\exp\left(-\frac{\sum_1^N i(x_i-\mu)^2}{2\sigma^2}\right)\right]\left(-\frac{\sum i(x_i-\mu)^2}{\cancel{2}\sigma^3}(-\cancel{2})\right)$$

$$= \frac{\exp\left(-\frac{\sum_1^N i(x_i-\mu)^2}{2\sigma^2}\right)}{\sigma^{N+3}}\left\{-N\cdot\sigma^2 + \sum_1^N i(x_i-\mu)^2\right\}$$

$$\boxed{\sigma = \sqrt{\frac{\sum_1^N i(x_i-\mu)^2}{N}}}$$

Se non conosco μ, lo devo stimare anch'esso dalle N misure tramite il valore medio $\langle x\rangle = \sum_1^N {}_i x_i / N$

$$\rightarrow \sigma = \sqrt{\frac{1}{N}\sum_1^N {}_i (x_i - \mu)^2} \geq \sqrt{\frac{1}{N}\sum_1^N {}_i (x_i - \langle x\rangle)^2}$$

σ viene sottostimato poiché $\sum_{i=1}^N (x_i-\mu)^2$ è minimo per $\mu = \langle x\rangle$!

La migliore approssimazione sarà allora:

$$\sigma = \sqrt{\frac{1}{N}\sum_1^N {}_i (x_i - \langle x\rangle)^2} \cdot \sqrt{\frac{N}{N-1}} = \sqrt{\frac{\sum_1^N {}_i (x_i - \langle x\rangle)^2}{N-1}}$$

Infatti: suppongo di avere M gruppi di misure, ciascuno formato da N misure ripetute

$$\{x_{ik}\} \qquad \begin{array}{l} i = 1, 2, \ldots, N \\ k = 1, 2, \ldots, M \end{array}$$

per semplicità assumo che ogni gruppo di N misure abbia lo stesso $\sigma = \sqrt{\sum_1^N {}_i (x_{ik} - \mu)^2 / N} \quad \forall k = 1, 2, \ldots, M$

$$\left[(x_{ik}-\mu)\right]^2 = \left[(x_{ik} - \langle x\rangle_k) + (\langle x\rangle_k - \mu)\right]^2 =$$

$$= (x_{ik} - \langle x\rangle_k)^2 + (\langle x\rangle_k - \mu)^2 + 2(x_{ik} - \langle x\rangle_k)(\langle x\rangle_k - \mu)$$

$$\underbrace{\frac{\sum_1^N {}_i (x_{ik}-\mu)^2}{N}}_{\sigma_x^2} = \frac{\sum_1^N {}_i (x_{ik} - \langle x\rangle_k)^2}{N} + \underbrace{\frac{\sum_1^N {}_i (\langle x\rangle_k - \mu)^2}{N}}_{\sigma_{\langle x\rangle}^2 = \frac{\sigma_x^2}{N}} + 0$$

Infatti, utilizzando la "propagazione quadratica" per l'incertezza:

$$\sigma_{<x>}^2 = \sum_{i=1}^N \left(\frac{\partial < x >_k}{\partial x_i}\right)^2 \sigma_x^2 = \sum_{i=1}^N \left(\frac{\partial}{\partial x_i}\left(\frac{\sum_{i=1}^N x_{ik}}{N}\right)^2 \sigma_x^2\right) = \sum_{i=1}^N \left(\frac{1}{N}\right)^2 \sigma_x^2 = N\frac{1}{N^2}\sigma_x^2 = \frac{\sigma_x^2}{N}$$

$$\sigma_x^2 = \frac{\sum_1^N{}_i \left(x_{ik} - \langle x \rangle_k\right)^2}{N} + \frac{\sigma_x^2}{N}$$

$$\sigma_x^2 \underbrace{\left(1 - \frac{1}{N}\right)}_{\frac{N-1}{N}} = \frac{\sum_1^N{}_i \left(x_{ik} - \langle x \rangle_k\right)^2}{N}$$

$$\boxed{\sigma_x = \sqrt{\frac{\sum_1^N{}_i \left(x_{ik} - \langle x \rangle_k\right)^2}{N-1}}} \Leftarrow$$

mentre :

$$\boxed{\sigma_{\langle x \rangle} = \frac{\sigma_x}{\sqrt{N}}} = \sqrt{\frac{\sum_1^N{}_i \left(x_{ik} - \langle x \rangle_k\right)^2}{(N-1)\cdot N}} \Leftarrow$$

Osservazione :

$$\frac{1}{N} \sum_1^N{}_i \left(x_{ik} - \langle x \rangle_k\right)\left(\langle x \rangle_k - \mu\right) = \frac{1}{N}\sum_1^N{}_i \left[x_{ik}\langle x \rangle_k - \mu\, x_{ik} - \langle x \rangle_k^2 + \langle x \rangle_k \cdot \mu\right] =$$

$$= \langle x \rangle_k \underbrace{\boxed{\frac{\sum_1^N{}_i x_{ik}}{N}}}_{\langle x \rangle_k} - \mu \underbrace{\boxed{\frac{\sum_1^N{}_i x_{ik}}{N}}}_{\langle x \rangle_k} - \langle x \rangle_k^2 \frac{\cancel{N}}{\cancel{N}} + \langle x \rangle_k \mu \frac{\cancel{N}}{\cancel{N}} =$$

$$= \langle x \rangle_k^2 - \mu \langle x \rangle_k - \langle x \rangle_k^2 + \langle x \rangle_k \cdot \mu = 0$$

$$\sigma(a,b) = \frac{1}{N-1} \sum (a - \bar{a})(b - \bar{b})$$

$$\sum_{1\,k}^{N} (a_k - \bar{a})(b_k - \bar{b}) = \sum_{1\,k}^{N} \left(a_k b_k - \bar{b}\, a_k - \bar{a}\, b_k + \bar{a}\,\bar{b} \right) =$$

$$= N \frac{\sum_{1}^{N} k (a_k\, b_k)}{N} - N\bar{b} \frac{\sum_{1\,k}^{N} a_k}{N} - N\bar{a} \frac{\sum_{1}^{N} k\, b_k}{N} + \sum_{1}^{N} k (\bar{a}\,\bar{b}) =$$

$$= N\langle ab\rangle - N\langle b\rangle\langle a\rangle - N\langle a\rangle\langle b\rangle + N\langle a\rangle\langle b\rangle =$$

$$= N\left[\langle ab\rangle - 2\langle a\rangle\langle b\rangle + \langle a\rangle\langle b\rangle \right] = N\left[\langle ab\rangle - \langle a\rangle\langle b\rangle \right]$$

$$Y = a x_1 + b x_2$$

$$\sigma^2(Y) = a_1^2 \sigma^2(x_1) + a_2^2 \sigma^2(x_2) + 2 a_1 a_2\, \sigma(x_1, x_2)$$

Limiti della media aritmetica

- G.F. x misurata più volte in "laboratori" diversi, cioè con precisioni diverse:

↳ $x_{11}, x_{12}, x_{13}, \ldots\ldots, x_{1N_1}$

$$\langle x_1 \rangle = \sum_{1}^{N_1} {}_i \; x_{1i} / N_1$$

$$\sigma(\langle x_1 \rangle) = \sqrt{\frac{\sum_{1}^{N_1} {}_i (x_{1i} - \langle x_1 \rangle)^2}{(N_1 - 1) \cdot N_1}}$$

$$\boxed{\langle x_1 \rangle \pm \sigma(\langle x_1 \rangle)}$$

↳ $x_{21}, x_{22}, x_{23}, \ldots\ldots, x_{2N_2}$

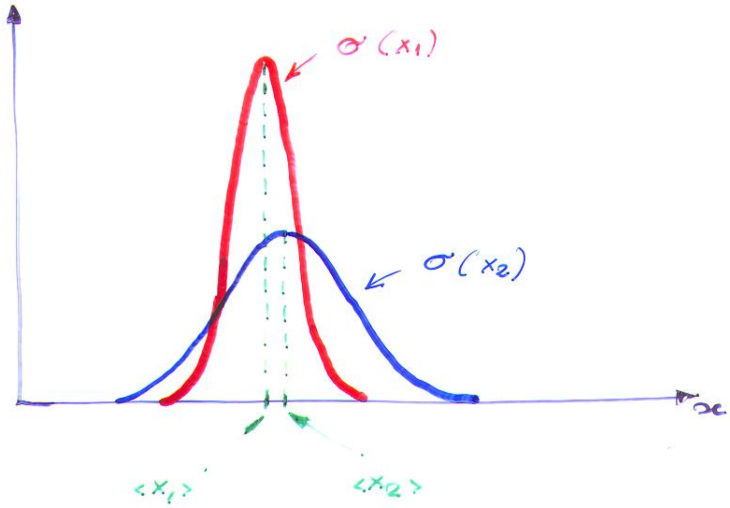
$$\langle x_2 \rangle = \sum_{1}^{N_2} {}_i \; x_{2i} / N_2$$

$$\sigma(\langle x_2 \rangle) = \sqrt{\frac{\sum_{1}^{N_2} {}_i (x_{2i} - \langle x_2 \rangle)^2}{(N_2 - 1) \cdot N_2}}$$

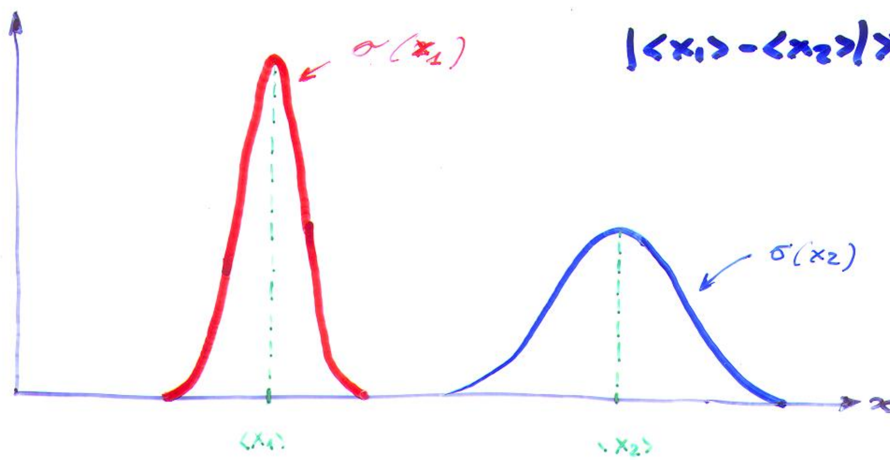
$$\boxed{\langle x_2 \rangle \pm \sigma(\langle x_2 \rangle)}$$

come possiamo ottenere una unica stima della G.F. x ?

→ misure consistenti : $|\langle x_1\rangle - \langle x_2\rangle| \lessapprox \sigma_{\langle x_1\rangle}, \sigma_{\langle x_2\rangle}$



$|\langle x_1\rangle - \langle x_2\rangle| \gg \sigma_{\langle x_1\rangle}, \sigma_{\langle x_2\rangle}$



.... e nei casi intermedi ?.....

→ Misure inconsistenti ?
Forse una delle serie di misure o entrambe sono affette da errore sistematico !

Media pesata: combinazione di misure aventi diversa precisione

→ Le misure provengono da distribuzioni limite tutte di tipo gaussiano ma aventi diversi valori della deviazione standard σ_i

$$\{x_i, \sigma_i\} \quad \text{con} \quad i = 1, 2, 3, \ldots, N, \quad f(x_i) = \frac{1}{\sqrt{2\pi}\,\sigma_i} e^{-\frac{(x_i-\mu)^2}{2\sigma_i^2}}$$

→ Principio della massima verosimiglianza: Applicato per la migliore stima del valore vero incognito μ

$$P(x_1, x_2, \ldots, x_N) = P(x_1) \cdot P(x_2) \cdot \ldots \cdot P(x_N)$$

$$\propto \frac{1}{\sigma_1} e^{-\frac{(x_1-\mu)^2}{2\sigma_1^2}} \cdot \frac{1}{\sigma_2} e^{-\frac{(x_2-\mu)^2}{2\sigma_2^2}} \cdot \ldots \cdot \frac{1}{\sigma_N} e^{-\frac{(x_N-\mu)^2}{2\sigma_N^2}}$$

$$= \prod_{i=1}^{N} \frac{1}{\sigma_i} \exp\left(-\frac{(x_i-\mu)^2}{2\sigma_i^2}\right) = \frac{\exp\left(-\sum_1^N{}_i \frac{(x_i-\mu)^2}{2\sigma_i^2}\right)}{\sigma_1 \cdot \sigma_2 \cdot \ldots \cdot \sigma_N}$$

$$\boxed{\frac{\partial P}{\partial \mu} = 0} \quad 0 = \frac{\partial}{\partial \mu}\left\{\sum_1^N{}_i \frac{(x_i-\mu)^2}{2\sigma_i^2}\right\} = \sum_1^N{}_i \frac{2(x_i-\mu)(-1)}{2\sigma_i^2} =$$

$$= -\left[\sum_1^N{}_i \frac{x_i}{\sigma_i^2} - \sum_1^N{}_i \frac{\mu}{\sigma_i^2}\right]$$

$$\sum_1^N{}_i \frac{x_i}{\sigma_i^2} = \mu \sum_1^N{}_i \frac{1}{\sigma_i^2} \qquad \boxed{\mu = \sum_1^N{}_i \frac{x_i}{\sigma_i^2} \Big/ \sum_1^N{}_i \frac{1}{\sigma_i^2}}$$

Un buon estimatore per μ non è più la semplice media aritmetica, ma la media pesata con $\boxed{\text{pesi} = \frac{1}{\sigma_i^2} = \frac{1}{\text{varianza}}}$

$$f(x, y)$$

$$\Delta f = \frac{\partial f}{\partial x}\Delta x + \frac{\partial f}{\partial y}\Delta y$$

$$\sigma^2(f(x)) = \left(\frac{\partial f}{\partial x}\right)^2 \sigma^2(x) + \left(\frac{\partial f}{\partial y}\right)^2 \sigma^2(y)$$

$$\sigma^2(f(x)) = \sum_{k} \left(\frac{\partial f}{\partial x_k}\right)^2 \sigma^2(x_k)$$

→ Varianza della media pesata dalla propagazione delle incertezze statistiche:

$$\sigma^2_{\substack{media\\pesata}} = \sum_1^N{}_i \left(\left[\frac{\partial}{\partial x_i} \left(\frac{\sum_1^N{}_k \frac{x_k}{\sigma_k^2}}{\sum_1^N{}_k \frac{1}{\sigma_k^2}} \right) \right]^2 \cdot \sigma_i^2 \right) =$$

$$= \sum_1^N{}_i \left(\left[\frac{\frac{1}{\sigma_i^2}}{\sum_1^N{}_k \frac{1}{\sigma_k^2}} \right]^2 \sigma_i^2 \right) = \sum_1^N{}_i \left(\frac{\frac{1}{\sigma_i^2}}{\left(\sum_1^N{}_k \frac{1}{\sigma_k^2} \right)^2} \right) = \frac{1}{\sum_1^N{}_k \frac{1}{\sigma_k^2}}$$

$$\boxed{\frac{1}{\sigma^2_{\substack{media\\pesata}}} = \sum_1^N{}_\ell \frac{1}{\sigma_\ell^2}} \qquad \sigma_{\substack{media\\pesata}} = \frac{1}{\sqrt{\sum_\ell \frac{1}{\sigma_\ell^2}}}$$

→ naturalmente: ① media pesata = media aritmetica

se $\sigma_i = \sigma$ per ogni i

$$\mu = \frac{\sum_1^N{}_i \frac{x_i}{\sigma_i^2}}{\sum_1^N{}_i \frac{1}{\sigma_i^2}} = \frac{\sum_1^N{}_i \frac{x_i}{\sigma^2}}{\sum_1^N{}_i \frac{1}{\sigma^2}} = \frac{\frac{1}{\sigma^2} \sum_1^N{}_i x_i}{N \cdot \frac{1}{\sigma^2}} = \frac{\sum_1^N{}_i x_i}{N} = \langle x \rangle$$

② varianza della media pesata $= \frac{\sigma^2}{N}$ se $\sigma_i = \sigma$ per ogni i

$$\frac{1}{\sigma^2_{\substack{media\\pesata}}} = \sum_1^N{}_i \frac{1}{\sigma_i^2} = \sum_1^N{}_i \frac{1}{\sigma^2} = \frac{N}{\sigma^2}$$

Esempi numerici:

- $$c_1 = (299'774 \pm 2) \ \text{Km/s}$$
  $$c_2 = (299'778 \pm 4) \ \text{Km/s}$$
  qual'è il valore più attendibile ?

  $$|c_1 - c_2| = 4 \ \text{Km/s}$$

  $$\rightarrow \sigma_{\substack{\text{media}\\ \text{pesata}}} = \frac{1}{\sqrt{\left(\frac{1}{2}\right)^2 + \left(\frac{1}{4}\right)^2}} \cong 1,789 \ \text{Km/s}$$

  $$c_{\substack{\text{media}\\ \text{pesata}}} = \frac{\frac{299'774}{(2)^2} + \frac{299'778}{(4)^2}}{\left(\frac{1}{2}\right)^2 + \left(\frac{1}{4}\right)^2} \cong 299774.800 \ \text{Km/s}$$

  $$\Rightarrow (299'774.8 \pm 1.8) \ \text{Km/s}$$

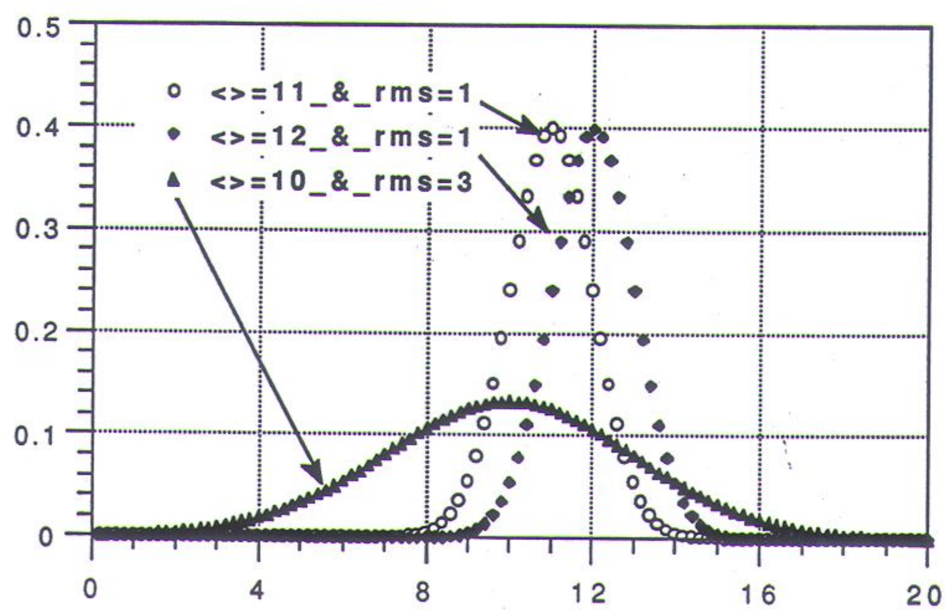
- $$R_1 = (11 \pm 1) \ \Omega$$
  $$R_2 = (12 \pm 1) \ \Omega$$
  $$R_3 = (10 \pm 3) \ \Omega$$

  $$|R_1 - R_3| = 1 \ \Omega$$
  $$|R_2 - R_3| = 2 \ \Omega$$
  $$|R_1 - R_2| = 1 \ \Omega$$

  $$\rightarrow \sigma_{\substack{\text{media}\\ \text{pesata}}} = \frac{1}{\sqrt{\left(\frac{1}{1}\right)^2 + \left(\frac{1}{1}\right)^2 + \left(\frac{1}{3}\right)^2}} \cong 0,688 \ \Omega$$

  $$R_{\substack{\text{media}\\ \text{pesata}}} = \frac{\frac{11}{(1)^2} + \frac{12}{(1)^2} + \frac{10}{(3)^2}}{\frac{1}{(1)^2} + \frac{1}{(1)^2} + \frac{1}{(3)^2}} \cong 11,421 \ \Omega$$

  $$\Rightarrow (11,42 \pm 0,69) \ \Omega$$

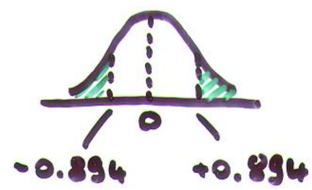


0.5
0.4
0.3
0.2
0.1
0
<>=11_&_rms=1
<>=12_&_rms=1
<>=10_&_rms=3
0
4
8
12
16
20

Prima di calcolare la Media Pesata va verificata la compatibilità:

- $$\frac{c_1 - c_2}{\sqrt{\sigma^2(c_1) + \sigma^2(c_2)}} = \frac{4}{\sqrt{4+16}} = 0.894$$

  $P(|t| \geq 0.894) = 37{,}1\%$



- $$\frac{R_1 - R_3}{\sqrt{\sigma^2(R_1) + \sigma^2(R_3)}} = \frac{1}{\sqrt{1+9}} = 0.316$$

  $P(|t| \geq 0.316) = 75{,}3\%$

  $$\frac{R_2 - R_3}{\sqrt{\sigma^2(R_2) + \sigma^2(R_3)}} = \frac{2}{\sqrt{1+9}} = 0.632$$

  $P(|t| \geq 0.632) = 52{,}5\%$

  $$\frac{R_1 - R_2}{\sqrt{\sigma^2(R_1) + \sigma^2(R_2)}} = \frac{1}{\sqrt{1+1}} = 0.707$$

  $P(|t| \geq 0.707) = 48{,}4\%$

Incertezza percentuale della Media Pesata:

- $$\frac{\sigma(c_1)}{c_1} = 6.67 \times 10^{-6}$$

  $$\frac{\sigma(c_2)}{c_2} = 1.33 \times 10^{-5}$$

  $$\frac{\sigma(\bar{c})}{\bar{c}} = 6.00 \times 10^{-6}$$

- $$\frac{\sigma(R_1)}{R_1} = 9.1\%$$

  $$\frac{\sigma(R_2)}{R_2} = 8.3\%$$

  $$\frac{\sigma(R_3)}{R_3} = 30\%$$

  $$\frac{\sigma(\bar{R})}{\bar{R}} = 6.0\%$$

… l'incertezza sulla media pesata, sia come valore assoluto che come valore relativo e' sempre piu' piccola del piu' piccolo dei contributi usati per il suo calcolo.