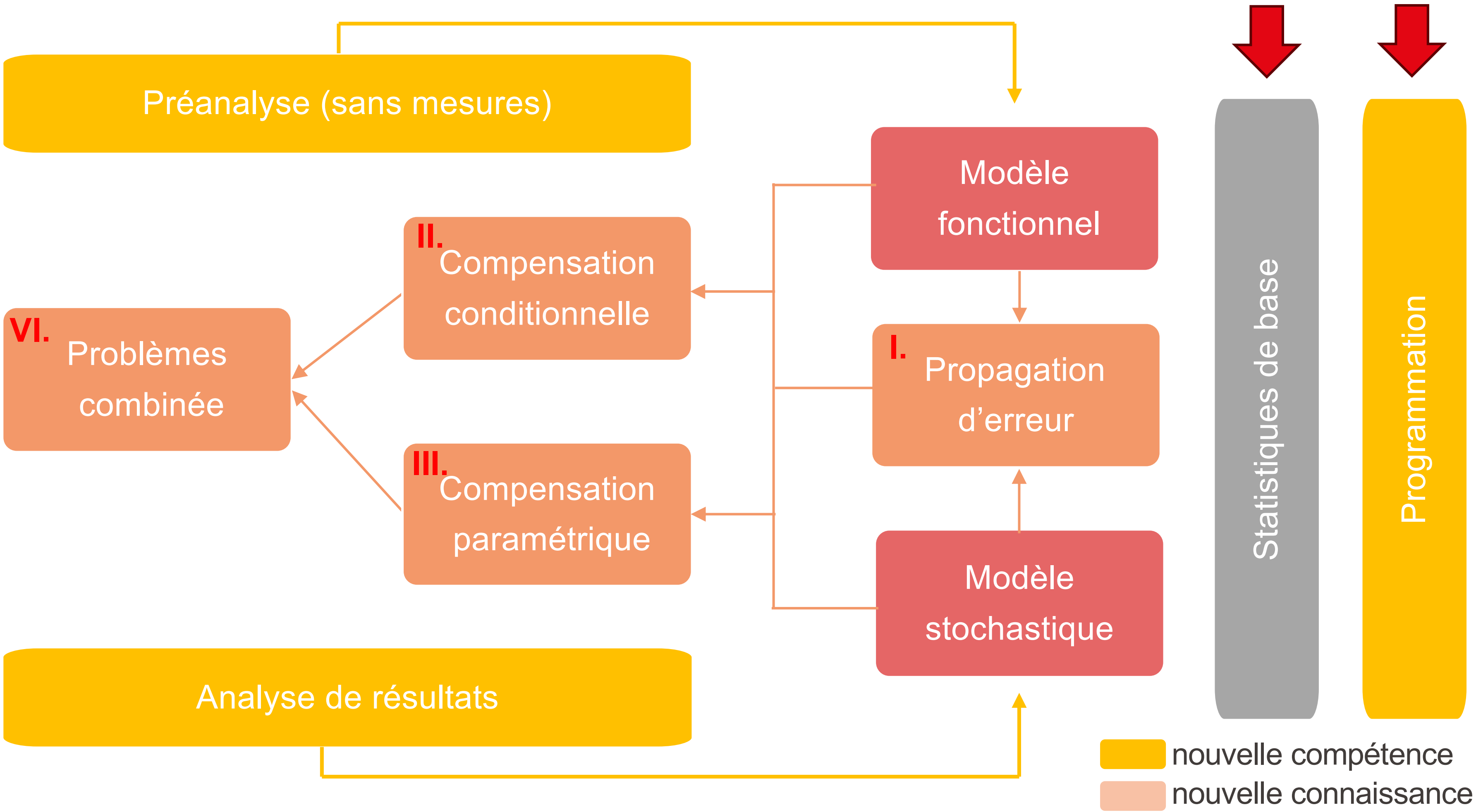
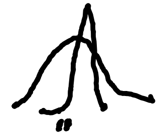
ME Cockpit de de matières et de compétences

- Comment réussir à observer et estimer avec confiance?



EPFL

Survol des éléments les plus utiles



- Théorème central limite (TCM): $\sum^{\infty} x_i \longrightarrow$

- L'approche fréquentiel: $F(x_i) \approx P(X < x_i)$

$$\varphi_x = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{x_i - E\{x\}}{2\sigma^2}}$$

- L'espérance: $E\{X\} = \frac{1}{n}\sum_{i=1}^{n} x_i \cdot p$ $n >> 1$ $n \to \infty$
- Variance: $E\{(x - E\{x\})^2\}$
- PDF (fon. de densité ~~de~~ de probabilité) de plus utiles associées?
  - Student T = test $T = [\hat{E}\{x\} - \hat{E}\{y\}]$? même moyenne variable
  - Chi^2 ~ T(sample var) $\Rightarrow$ distr. normal < oui / non
  - Ficher entre DEUX "sample" var $T\left[\frac{\text{var}1}{\text{var}2}\right]$ < même / pas de même

$\hat{\cdot} = \text{\hat}.\ l = \hat{l} \quad \bar{\cdot} = \text{\bar}\ l = \bar{l}$

Survol des éléments avec les nouveaux symboles

Statistiques de base

	Observation	STAT y_i	ME l_i	Formule $i=1 \dots n$	Note
Valeur	Vrais	μ	$\check{l}$ $(\check{\cdot})$		
	Estimé	$\hat{\mu}$ $(\hat{\cdot})$	$\hat{\cdot}$ $\bar{\cdot}$ moyenne		
La moyenne	Vrais + erreur	$\mu + r_i$	$\check{l} + \varepsilon_i$	$P(\check{l} \mid l_1 \dots l_n)$	
	Estimé + erreur (»sample »)	$\hat{y} + r_i$ résidue	$\hat{l} + v$	$P(\hat{l} \mid l_1 \dots l_n)$	v = Verbesserung
Variance	Vrais	$E[X^2] - (E[X])^2$	$\sigma^2 = \frac{\sum \varepsilon_i^2}{n}$		
	Estimé (»sample »)	$S = \frac{\sum r_i^2}{n-1}$	$\hat{\sigma}^2 = \frac{\sum v_i^2}{n-1}$		Attention !
L'écart-type		$\sqrt{S}$	$\sqrt{\hat{\sigma}^2}$		d'une observation y_i, l_i

Formules usuelles 2D

	erreurs vraies	erreurs résiduelles
	$n \longrightarrow \infty$	$n < \infty$
	définition	estimation
	$\ell_{x_i} - \varepsilon_{x_i} = \check{x}$	$\ell_{x_i} - \hat{v}_{x_i} = \hat{x}$
	$\ell_{y_i} - \varepsilon_{y_i} = \check{y}$	$\ell_{y_i} - \hat{v}_{y_i} = \hat{y}$
Variance de ℓ_{x_i}	$\boxed{\sigma_x^2 = \frac{\boldsymbol{\varepsilon}_x^\top \boldsymbol{\varepsilon}_x}{n}}$	$\boxed{\hat{\sigma}_x^2 = \frac{\hat{\boldsymbol{v}}_x^\top \hat{\boldsymbol{v}}_x}{n-1}}$
Variance de ℓ_{y_i}	$\boxed{\sigma_y^2 = \frac{\boldsymbol{\varepsilon}_y^\top \boldsymbol{\varepsilon}_y}{n}}$	$\boxed{\hat{\sigma}_y^2 = \frac{\hat{\boldsymbol{v}}_y^\top \hat{\boldsymbol{v}}_y}{n-1}}$
Covariance de ℓ_{x_i} et ℓ_{y_i}	$\boxed{\sigma_{xy} = \frac{\boldsymbol{\varepsilon}_x^\top \boldsymbol{\varepsilon}_y}{n}}$	$\boxed{\hat{\sigma}_{xy} = \frac{\hat{\boldsymbol{v}}_x^\top \hat{\boldsymbol{v}}_y}{n-1}}$
Coefficient de corrélation de ℓ_{x_i} et ℓ_{y_i}	$\boxed{\rho_{xy} = \frac{\sigma_{xy}}{\sigma_x \cdot \sigma_y}}$	$\boxed{\hat{\rho}_{xy} = \frac{\hat{\sigma}_{xy}}{\hat{\sigma}_x \cdot \hat{\sigma}_y}}$

La même variance?
Estimateur vs. estimateur non biaisé

$$V(X) = E\{(X - E\{X\})^2\} = E\{X^2\} - (E\{X\})^2$$

		Function	Estimateur / normalisation
Matlab	Variance	`var(x)`	non-biaisé (default) / n-1
	Écart-type	`std(x)`	non-biaisé (default) / n-1
Python	Variance	`numpy.var(x)`	biaisé (défaut) / n $\times \frac{n}{n-1}$
	Écart-type	`numpy.std(x)`	biaisé (défaut) / n $\times \frac{n}{n-1}$

B - changer on par. par défaut pour obtenir non-biaisé

A pour obtenir non-biaisé