

# Test d'Hypothèse

## 1. Hypothèse Nulle

$H_0$  : la variable est uniquement influencée par des causes  $\mathcal{C}$  de variation.

$H_A$  : la variable est influencée par des causes Spéciales en plus des causes  $\mathcal{C}$  de variation.

## 2. Valeur de probabilité

$p$ -val.  $\begin{cases} < 5\% \Rightarrow \text{conclure Rejet } H_0 \Leftrightarrow \text{Accepter } H_A \\ > 5\% \Rightarrow \text{conclure Echec Rejet } H_0 \\ \Rightarrow \text{on considère que } H_0 \text{ est vraie} \end{cases}$

## 3. Interpréter

X1 CHARGE : Discret

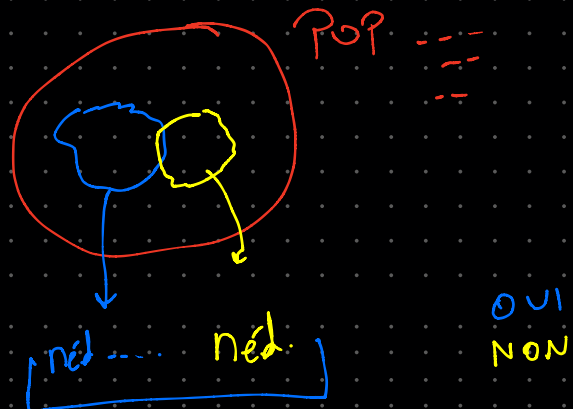
QUID si effet sur la tendance centrale?

1/ Test de la médiane de Mood  
Compare les médianes

$H_0$  : toutes les médianes sont égales = la charge n'a pas d'effet sur la médiane de la distance.

$P$ -val  $< 5\%$ . -->  $RH_0$ . -->  $AH_A$

La charge a un effet significatif sur la médiane de la distance

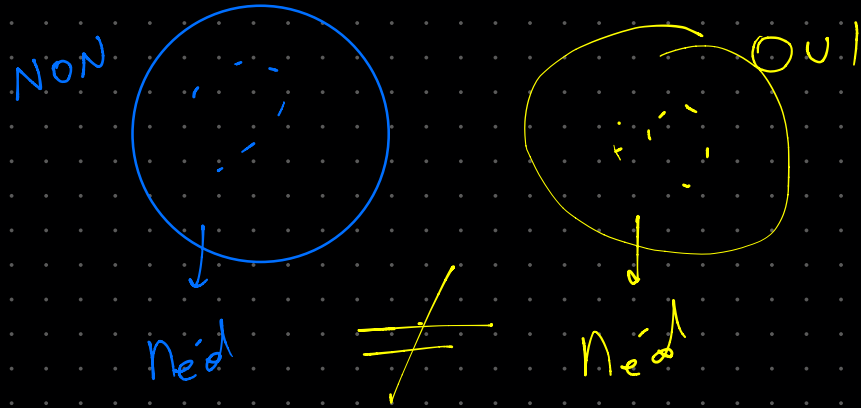


## Test

Hypothèse nulle  $H_0$  : les médianes de population sont toutes égales  
 Hypothèse alternative  $H_1$  : les médianes de population ne sont pas toutes égales

DL Khi deux Valeur de P

1 12.99 0.000

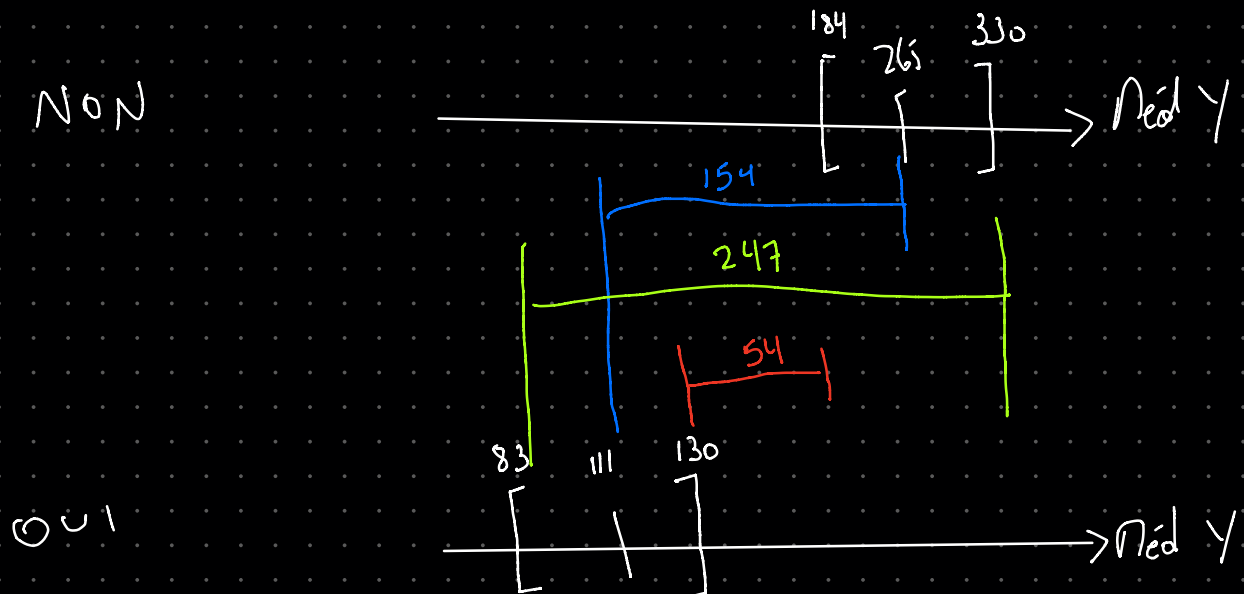


## Statistiques descriptives

Charge Médiane N  $\leq$  médiane globale N  $>$  médiane globale Q3 - Q1 IC médian à 95 %

NON	265	18	39	333.5	(184.332, 330)
OUI	111	57	35	137.5	(82.7149, 130)
Global	133				

95.0% IC pour médiane(NON) - médiane(OUI) : (65,227.57)



effet 154 [54 ; 247] IC 95%

X1 CHARGE : Discret

QUID si effet sur la variabilité?

2/ Test de l'égalité des variances

H0 : la charge n'a pas d'effet sur la variabilité de la distance.

Ou toutes les variances (écarts-types) sont égales.

HA : il y a un effet significatif

OU au moins une des variances est différente.

P-val < 5%. --> RH0 et AHA

Il y a une différence significative de variance en fonction de la charge

### Tests

Méthode	Statistique	
	du test	Valeur de P
Comparaisons multiples	—	0.000
Levene	36.18	0.000

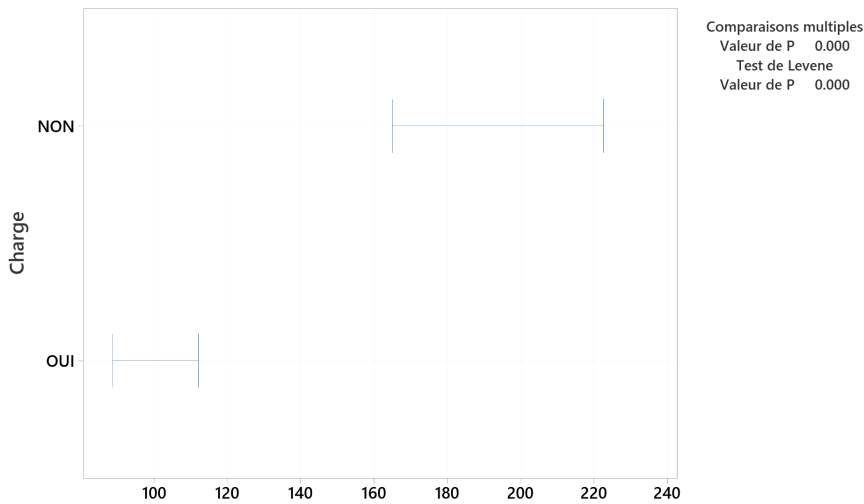
### Intervalle de confiance de Bonferroni à 95 % pour les écarts types

Charge	N	EcTyp	IC
NON	57	188.318	(160.465, 230.053)
OUI	92	98.707	(80.109, 124.660)

Niveau de confiance individuel = 97.5 %

### Test de l'égalité des variances : Distance mm en fonction de Charge

Intervalle de comparaisons multiples pour l'écart type,  $\alpha = 0.05$



Si les intervalles ne se superposent pas, les écarts types correspondants sont significativement différents.

X2 Hélico : Discret

Quid si hélico a un effet sur la TC de la distance?

1/ Test de la médiane de Mood

H0 : le X hélico n'a pas d'effet sur la médiane de la distance  
OU toutes les médianes sont égales

HA : Au moins une des médianes est différente.

P-val > 5% ERH0

On considère que H0 est vraie

On considère qu'il n'y a pas d'effet. On considère que toutes les médianes sont égales!

### Test

Hypothèse nulle

H<sub>0</sub> : les médianes de population sont toutes égales

Hypothèse alternative H<sub>1</sub> : les médianes de population ne sont pas toutes égales

DL Khi deux Valeur de P

3 2.43 0.489

Quid si Hélico a un effet sur la Var de la distance?

2/ Test de l'égalité des variances

H0 : tous les écarts types sont égaux

HA : au moins un des Ec T est différent

### Tests

Méthode	Statistique du test	Valeur de P
Comparaisons multiples		0.030
Levene	2.02	0.113

< 5% ERH0 => P'hélico a 1 effet signif. s/ éc.T. Dist  
simple - - - -

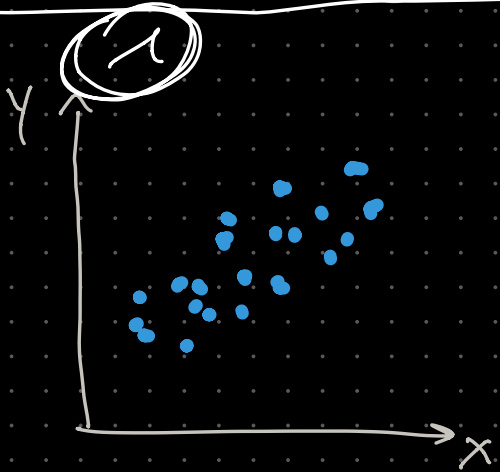
> 5% ERH0 => on considère que H0 est vraie  
- - - - -  
d'effet s/ éc.T. que l'hélico n'a pas  
on considère que tous les éc.t. =

Y CONTINU <math>\alpha</math> X CONTINU

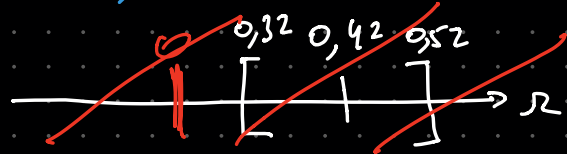
(1) check si corrélation / NON -> stop

OUI → tests (2)

## ② check régression



$$r = 0,42$$



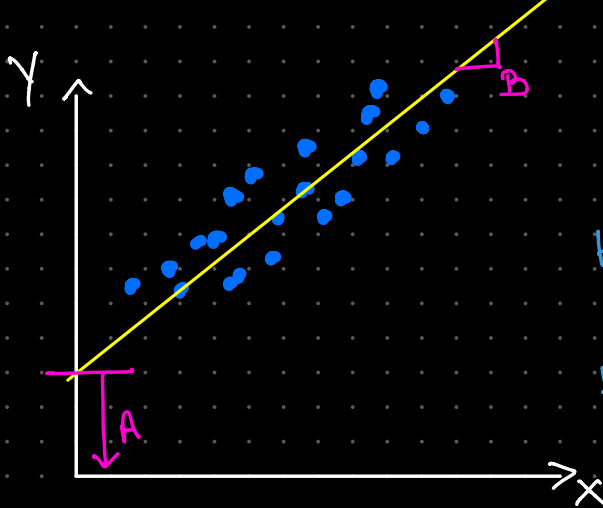
### Test Hypothèse

$H_0: r = 0 \Leftrightarrow X \text{ et } Y \text{ ne sont PAS corrélés.}$

$H_A: r \neq 0 \Leftrightarrow X \text{ et } Y \text{ présentent une corrélation SIGNIFIC.}$

②

## REGRESSION SIMPLE



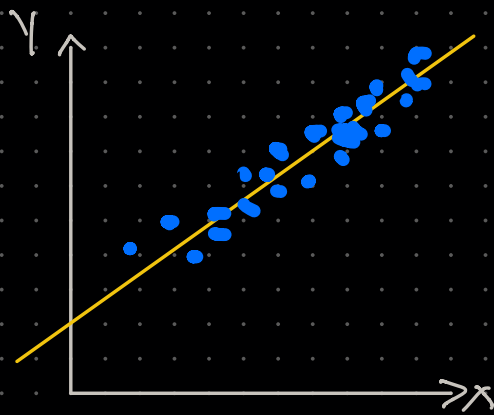
Modèle de régression

$$Y = A + B \cdot X$$

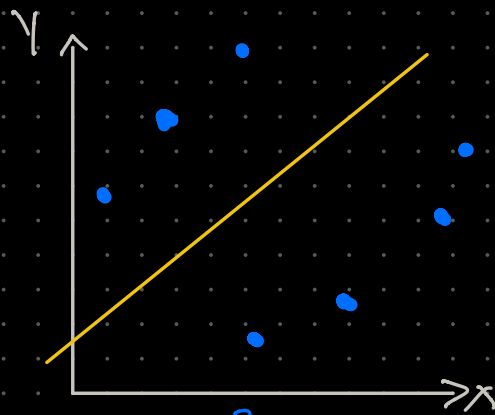
$H_0: \text{de modèle n'existe pas}$   
ou " $B=0$ "

$H_A: \text{de modèle existe}$

Si le modèle existe, check valeur  $R^2$



$R^2 \uparrow \uparrow$



$R^2 \downarrow$

$R^2$

Coefficient de régression

indique la proportion du comportement des données qui est expliquée par le modèle

•  $0 \leq R^2 \leq 100\%$

$= (r)^2$

REGRESSION LINEAIRE MULTIPLE

$$Y = A + B_1 \cdot X_1 + B_2 \cdot X_2 + B_3 \cdot X_3$$

① Test d'Hypoth. pour valider que le modèle complet EXISTE

$H_1$ : le modèle n'existe pas

② Tests d'Hypothèse pour valider l'effet de chaque X

$$Y = A + \underbrace{\beta_1 \cdot X_1}_{\downarrow \text{Ho: } \beta_1 = 0 \text{ p-val...}} + \underbrace{\beta_2 \cdot X_2}_{\downarrow \text{Ho: } \beta_2 = 0 \text{ p-val...}} + \underbrace{\beta_3 \cdot X_3}_{\downarrow \text{Ho: } \beta_3 = 0 \text{ p-val...}}$$

On ne gardera dans le modèle que les Xs qui ont un effet significatif

$\downarrow$   
 Coeff signif  $\neq 0$   
 $\Leftrightarrow$   
 p-val  $< 5\%$

X3 Hauteur : continu

Quid s'il y a corrélation entre X hauteur et Y distance?

1/ Test de corrélation

H0 : r = 0  
Ha : r  $\neq$  0

### Corrélations de Pearson deux à deux

Echantillon 1	Echantillon 2	N	Corrélation	IC à 95% pour p	Valeur de P
Distance mm	Hauteur mm	149	0.061	(-0.101, 0.219)	0.462

P-val = 0,462 --> ERH0 on considère que H0 est vraie ... on considère qu'il n'y a pas de corrélation entre X et Y

STOP

ON IMAGINE QU'IL Y AIT UNE CORRELATION SIGNIFICATIVE!!!

2/ Régression

2.1/ Valider que le modèle existe :

H0 : il n'y a pas de modèle

P-val = 0,462 ERH0 , on considère que le modèle n'existe pas

#### Analyse de la variance

Source	Somme des			
	DL	carrés	CM	F
Régression	1	13498	13497.6	0.54
Erreur	147	3652773	24848.8	0.462
Total	148	3666271		

Si le modèle avait existé, ajuste-t-il bien les données ?

$R^2 = 0,37\%$

Le modèle explique 0,37. % du comportement des données. :(

#### Récapitulatif du modèle

R carré		
S	R carré	(ajust)
157.635	0.37%	0.00%

## EXERCICE VISCOSITE

Régression linéaire multiple

1. Validation du modèle

H0 : le modèle n'existe pas

P-val <5% RH0 , le modèle existe

#### Analyse de la variance

Source	DL	SomCar ajust	CM ajust	Valeur F	Valeur de p
Régression	4	3271.18	817.80	11.93	0.000
Délustrant	1	539.76	539.76	7.87	0.016
Débit	1	28.97	28.97	0.42	0.528
Température de trempe	1	56.17	56.17	0.82	0.383
Rapidité	1	1015.10	1015.10	14.80	0.002
Erreur	12	822.94	68.58		
Total	16	4094.12			

2. Validation des termes

Pour chaque terme

H0 : le facteur n'a pas d'effet sur Y

Délustrant : p-val <5% RH0, le facteur a un effet significatif sur Y --> Je le garde

Débit : p-val > 5% --> ERH0 , le facteur n'a pas d'effet significatif sur Y, on considère qu'il n'en a pas --> Je l'élimine du modèle

#### Coefficients

Terme	Coef	Coef ErT	Valeur de T	Valeur de p	FIV
Constante	-208	595	-0.35	0.733	
Délustrant	1201	428	2.81	0.016	1.77
Débit	184	283	0.65	0.528	2.83
Température de trempe	0.505	0.558	0.91	0.383	1.73
Rapidité	58.4	15.2	3.85	0.002	1.98





Lignes : Employés Colonne : Colonnes de la feuille de trava

	Joe	Mary	Sid	Donna	Louise	Total
#défauts	6	7	8	9	5	35
	8.07	5.24	4.51	9.59	7.60	
	0.53046	0.59152	2.70931	0.03610	0.88795	
#non-défauts	148	93	78	174	140	633
	145.93	94.76	81.49	173.41	137.40	
	0.02933	0.03271	0.14980	0.00200	0.04910	
Total	154	100	86	183	145	668

Contenu de la cellule

Dénombrement

Dénombrement attendu

Contribution au Khi deux

H0 : toutes les proportions sont égales


P-val > 5% ERH0 ==> Toutes les proportions sont considérées comme égales!

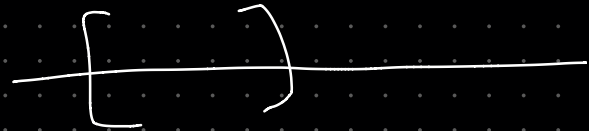
Test du Khi deux


	Khi deux	DL	Valeur de P
Pearson	5.018	4	0.295
Rapport de vraisemblance	4.627	4	0.328

1 cellules avec des dénombrements attendus inférieurs à 5.

*NA < 5%*

BN 

SU 

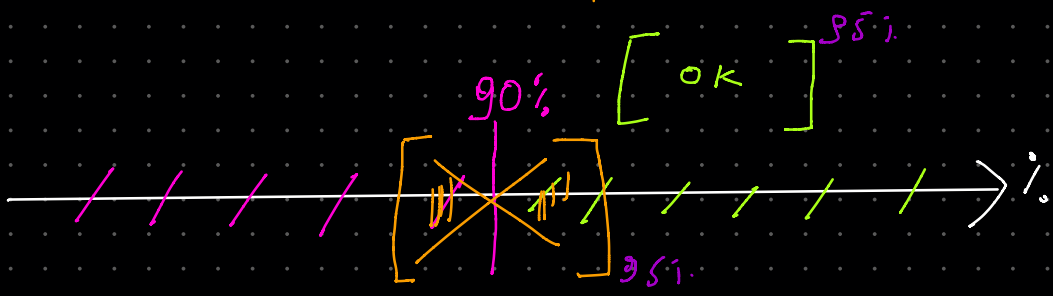
SP 

BG 

Р-удзси. R11.

	$P_{\text{нэп.}}$	$P_{\text{нэпро}}$	Ex
1) <u>I.C.</u>	—	—	—
2) KAPPA	—	—	—

Test H  
 3) KENDALL  
 Test H



[ NOK ]<sub>95%</sub>

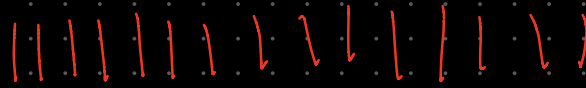
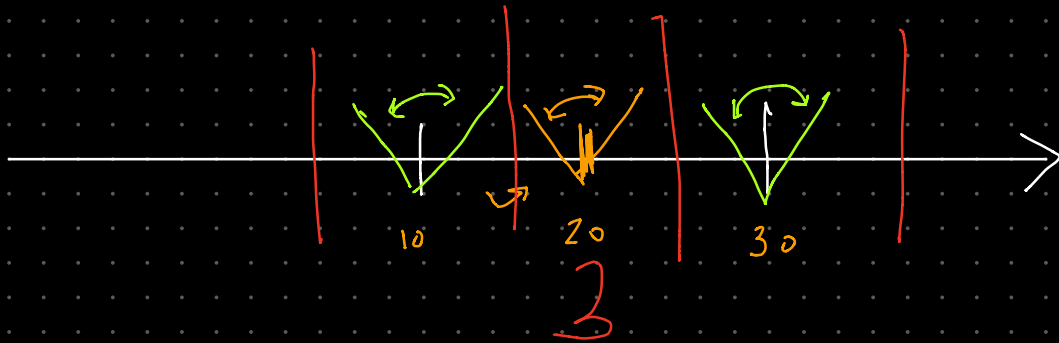
$$K = \frac{0,6 - 0,2}{1 - 0,2} = \frac{0,4}{0,8} = 0,5$$

$$0,8 = \frac{x - 0,2}{1 - 0,2}$$

$$0,8(1 - 0,2) + 0,2 = 0,82$$







# catégories distinctes

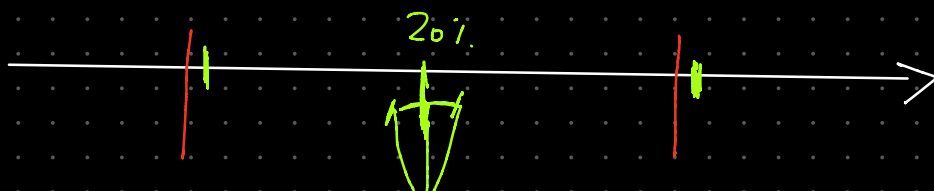
$H_0$ : pas répété

ANOVA

$H_0$ : toutes les moyennes se égales

Reproductible

$\neq$



1) Qu'est-ce qu'une base linéaire ?

→  $H_0$ : Base NON LINÉAIRE  
(pente nulle)  
→ p-vd.

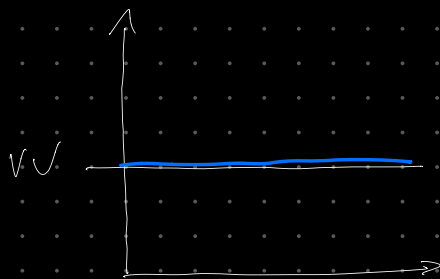
Si base constante

→ check base global (moyen)

Si base LIN.

→ check base locaux

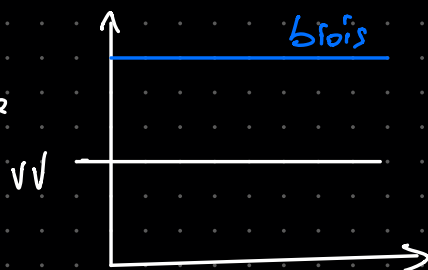
Exact



T. hypoth s/pente:  $E R H_0$

T. hypoth s/vol. base:  $E R H_0$

Inexact  
(Non Linéaire)

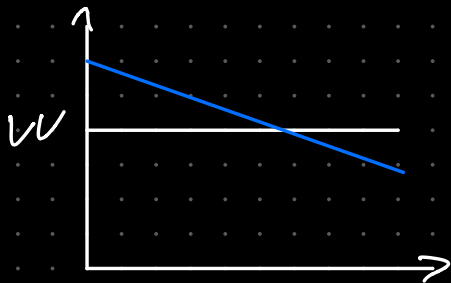


T. hypoth s/pente:  $E R H_0$

T. hypoth s/vol. base:  $R H_0$

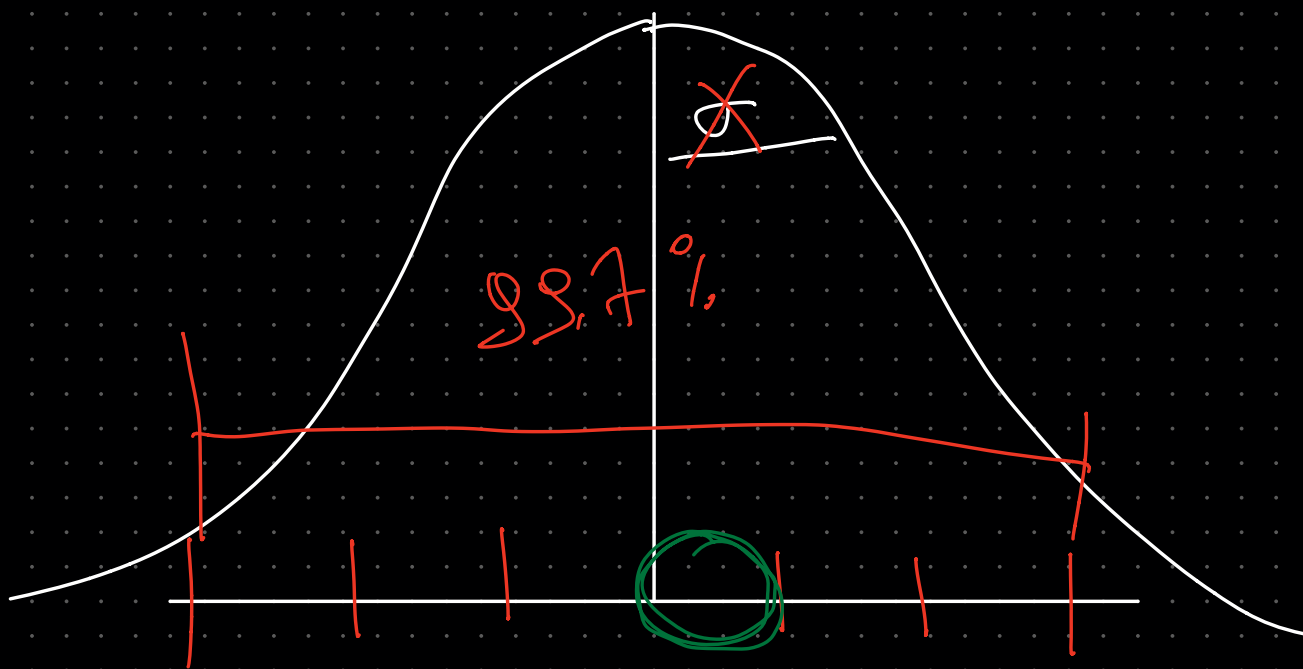


Interact  
Lineaire



T. hypoth s/ pente: R/0

T. hypoth s/ vol. bias: R/0

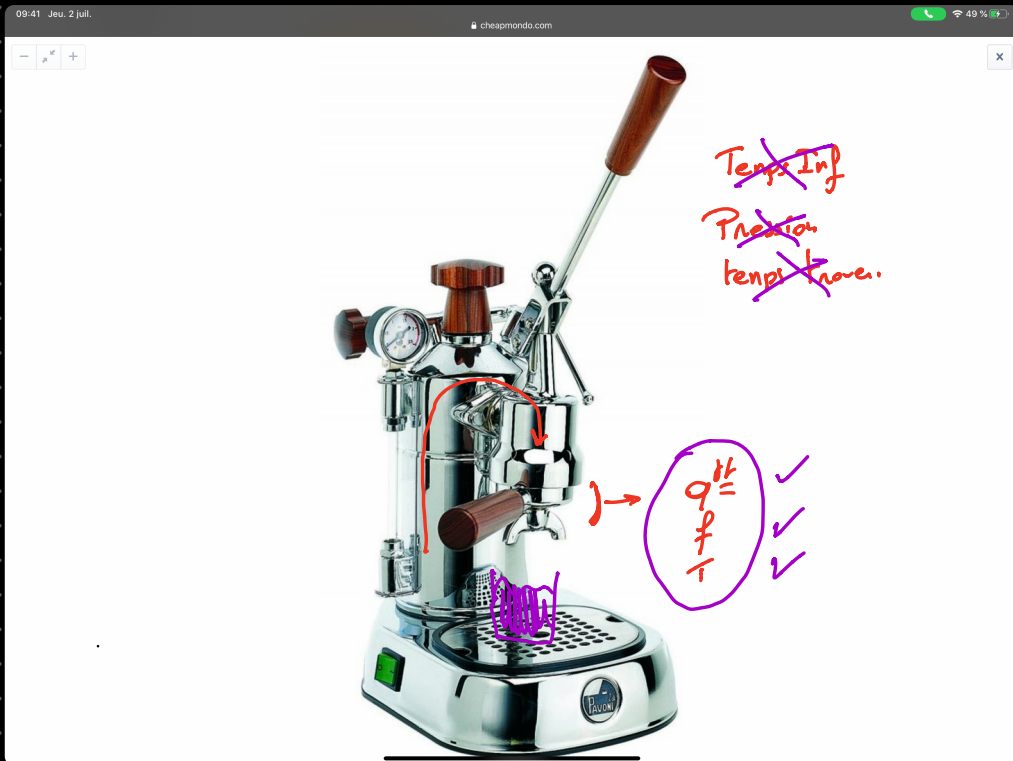


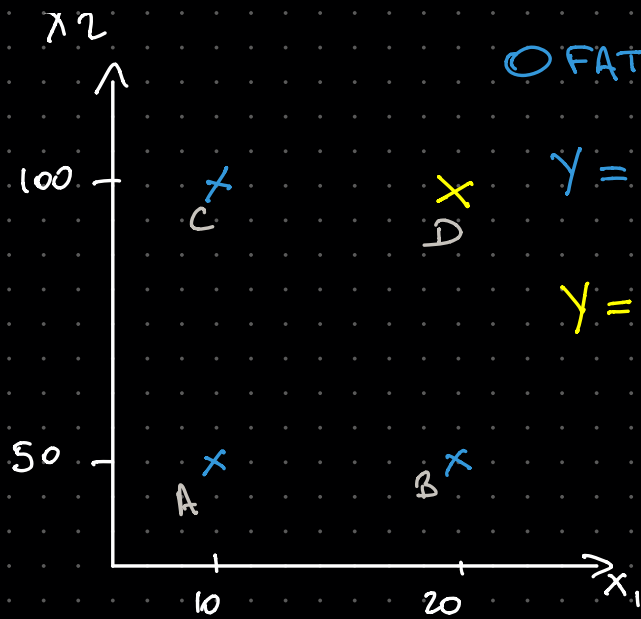
$$\frac{R}{6}$$

$$\% VE = \frac{\sigma}{n} = \frac{R/6}{1}$$

$\sigma$

$K/B$





O.F.A.T. → One Factor at A Time

$$Y = A + \beta_1 \cdot X_1 + \beta_2 \cdot X_2$$

effets principaux

$$Y = A + \beta_1 \cdot X_1 + \beta_2 \cdot X_2$$

$$+ \beta_{12} \cdot X_1 \cdot X_2$$

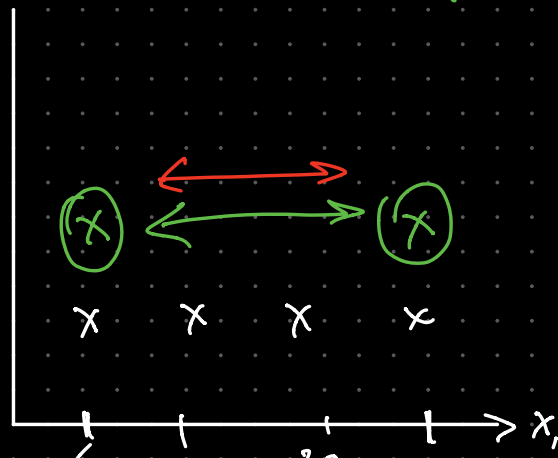
effet interaction

Point support : les combinaisons de conditions que l'on va tester

→ 4 points support

Essais : le nombre d'expériences à réaliser  
 si 4 fois chaque P.S. → 16 essais

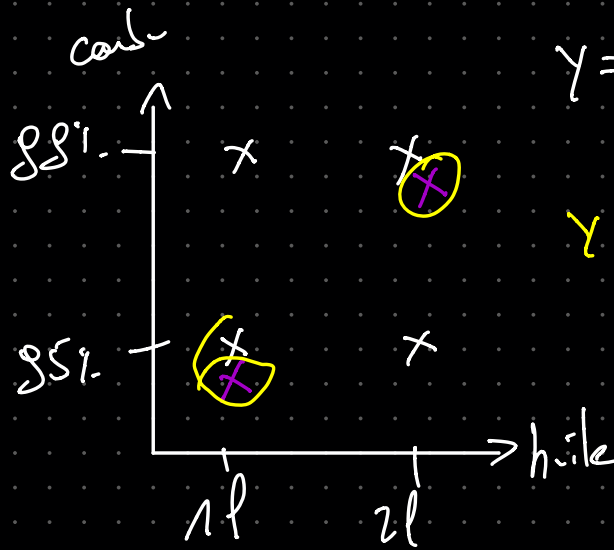
$$Y = \dots \beta_1 x_1 + \dots$$



5 10 ~ 30

$X_1$ : huile

$X_2$ : combu



$$Y = A + B_1 \cdot H + B_2 \cdot C$$

$$+ B_2 \cdot H \cdot C = 0,9$$

$$Y = A + B_1 \cdot \text{changement } "X_1"$$

$\swarrow$                        $\searrow$   
 H                                  C

effets principaux

Available Factorial Designs (with Resolution)

Runs	Factors														
	2	3	4	5	6	7	8	9	10	11	12	13	14	15	
4	Full	III													
8		Full	IV	<del>III</del>	<del>III</del>										
16			Full	V	<del>IV</del>	<del>IV</del>	<del>IV</del>	III	III	III	III	III	III	III	
32				Full	VI	<del>V</del>	<del>V</del>	IV	IV	IV	IV	IV	IV	IV	
64					Full	II	V	IV	IV	IV	IV	IV	IV	IV	
128						Full	III	VI	V	V	IV	IV	IV	IV	

" $x_1$ "      " $x_2$ "

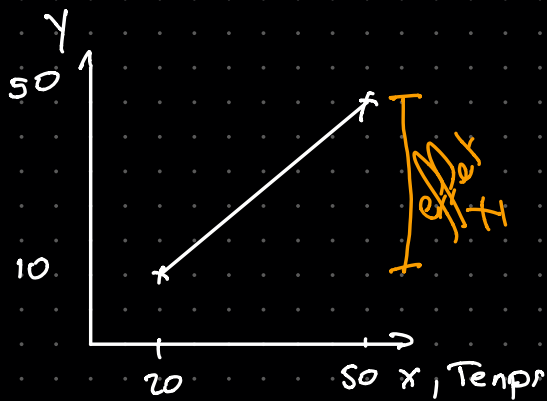
$$y = x_1 \quad x_2 \quad x_3 \quad \text{I}$$

$$x_1 \cdot x_2 \quad x_2 \cdot x_3 \quad x_1 \cdot x_3 \quad \text{II}$$

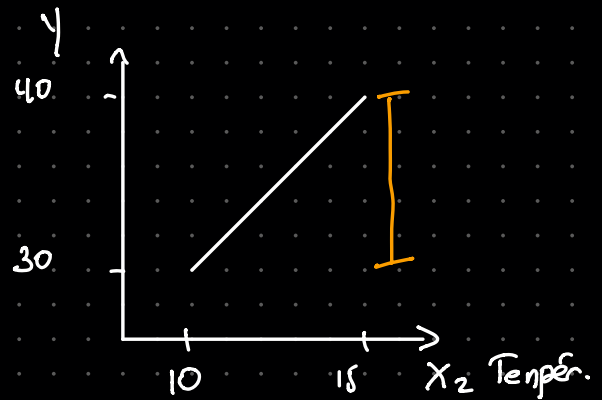
$$x_1 \cdot x_2 \cdot x_3 \quad \text{III}$$

~~IV~~

$$\text{Rés } 4 = 1 + 3 = 2 + 2$$



$$\text{effet } x_1 = 40$$



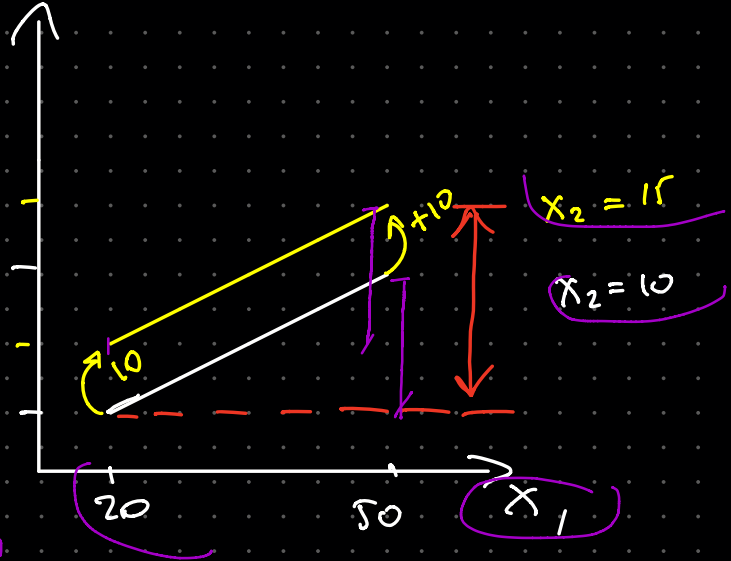
$$\text{effet } x_2 = 10$$

Pos d'interaction

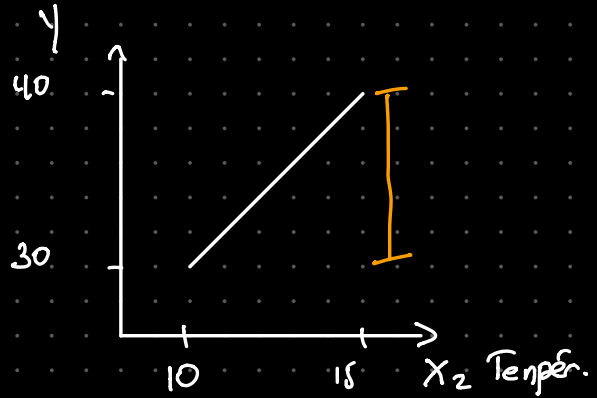
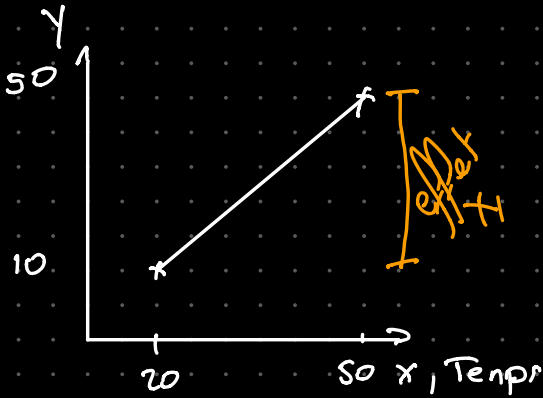
effet total

=  $\sum$  effets individ-els

$50 = 40 + 10$   
 $x_1 \quad x_2$



Interaction le1 x1 et x2

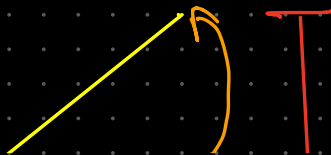


effet  $x_1 = 40$

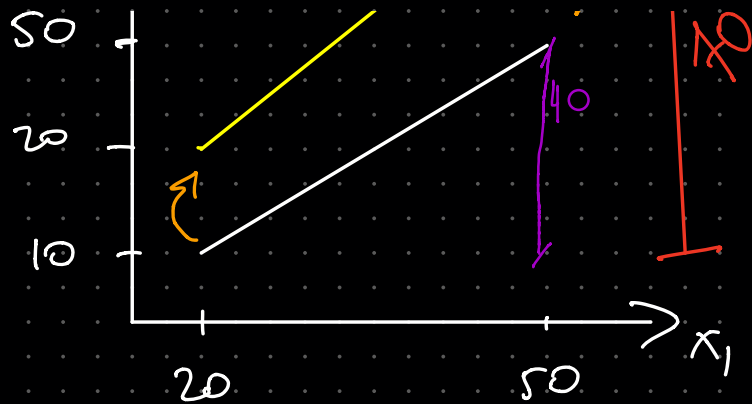
effet  $x_2 = 10$

80

$x_2 = 15$   
 $x_2 = 10$



effet  $\neq \sum$  effets indiv.



$$\text{Effet tot} = 70 = \sum \text{effets ind.} + \text{effet Inter.}$$

$$70 = 40 + 10 + 20$$

$$x_1 \quad x_2 \quad x_1 \cdot x_2$$

$$Y = A + b_1 x_1 + b_2 x_2 + b_3 x_3 \quad \text{I}$$

$$x_1 \cdot x_2 \quad x_2 \cdot x_3 \quad x_1 \cdot x_3 \quad \text{II}$$

$$x_1 \cdot x_2 \cdot x_3 \quad \text{III}$$

$$\text{IV}$$

## Jour 0

- 1) Quels sont nos facteurs ( $x_s$ )?
- 2) Quel domaine pour chaque facteur  
(min? max?)
- 3) { Suspecte interactions?  
est-ce que je veux les estimer?
- 4) Quid budget?  $\Leftrightarrow$  Quid #  
REPLICATIONS?

Objectif : maximiser répétitions

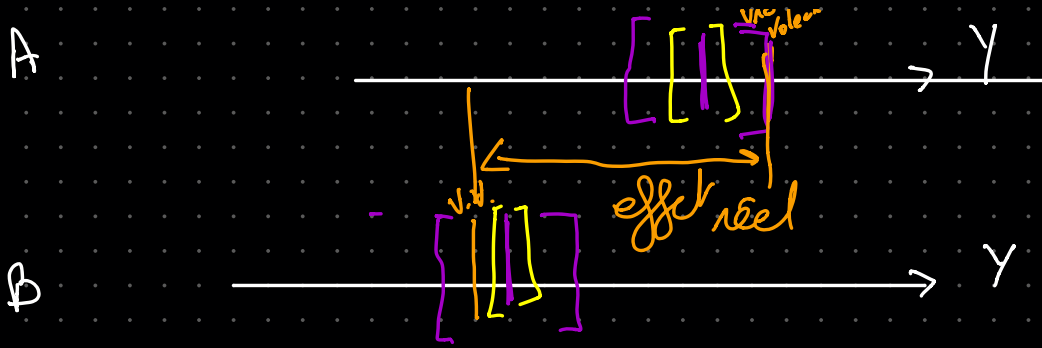
Retour à 13h

X

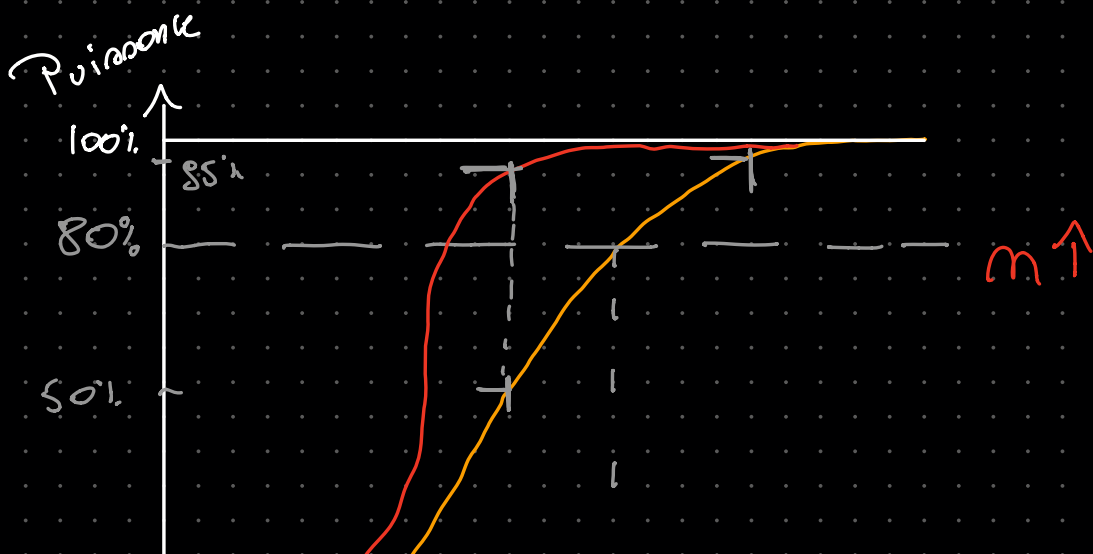
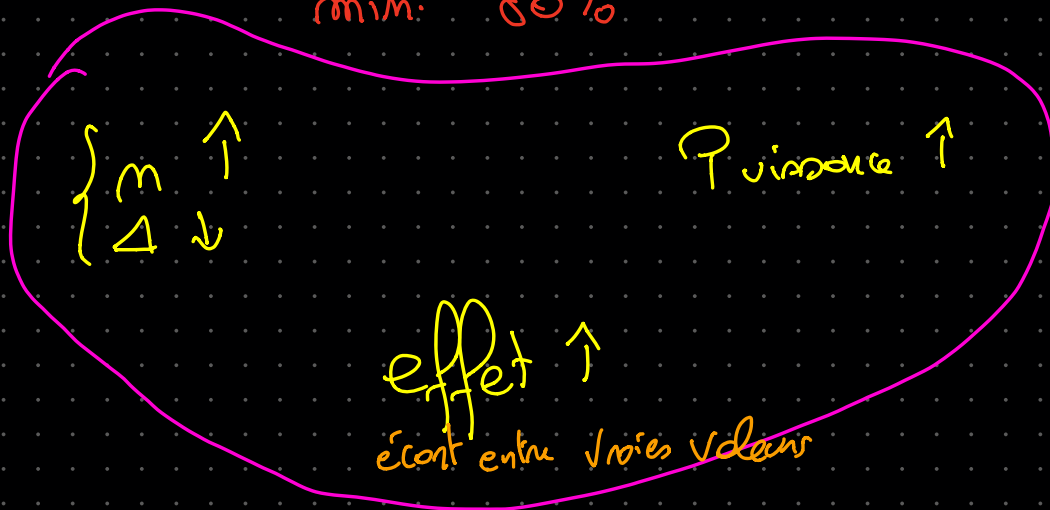
$\Delta = 2$  points

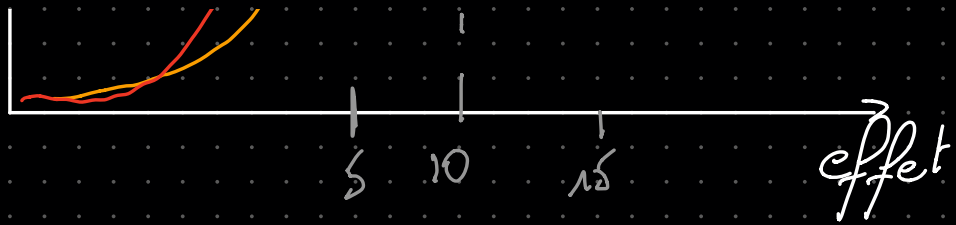
1.34





Puissance = probabilité de détecter un effet lorsqu'il existe  
 min. 80%

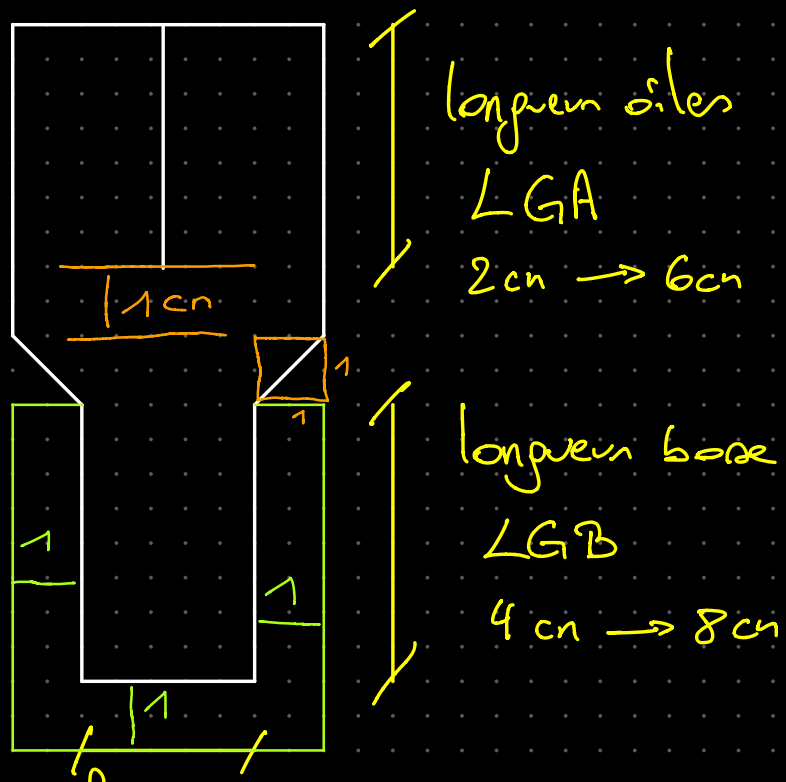




1) quid amples de l'effet à détecter

2) Puissance @ 80%

$n \propto \# \text{ Réplifications.}$



largeur  
Pb bosse  
1cm  $\rightarrow$  3cm

Effets d'interactions négligés

Détecter un plan d'effet de

40 nm.

Econt. Type  $\gamma = 38$  nm

Budget  $\approx 16$  essais

4 points support  
4 répétitions

Factoriel Fractionnaire  
3 facteurs  
F.F.  $\phi$  interact.  
répétitions

$$N = 2 \times 4 = 16$$

15h05

