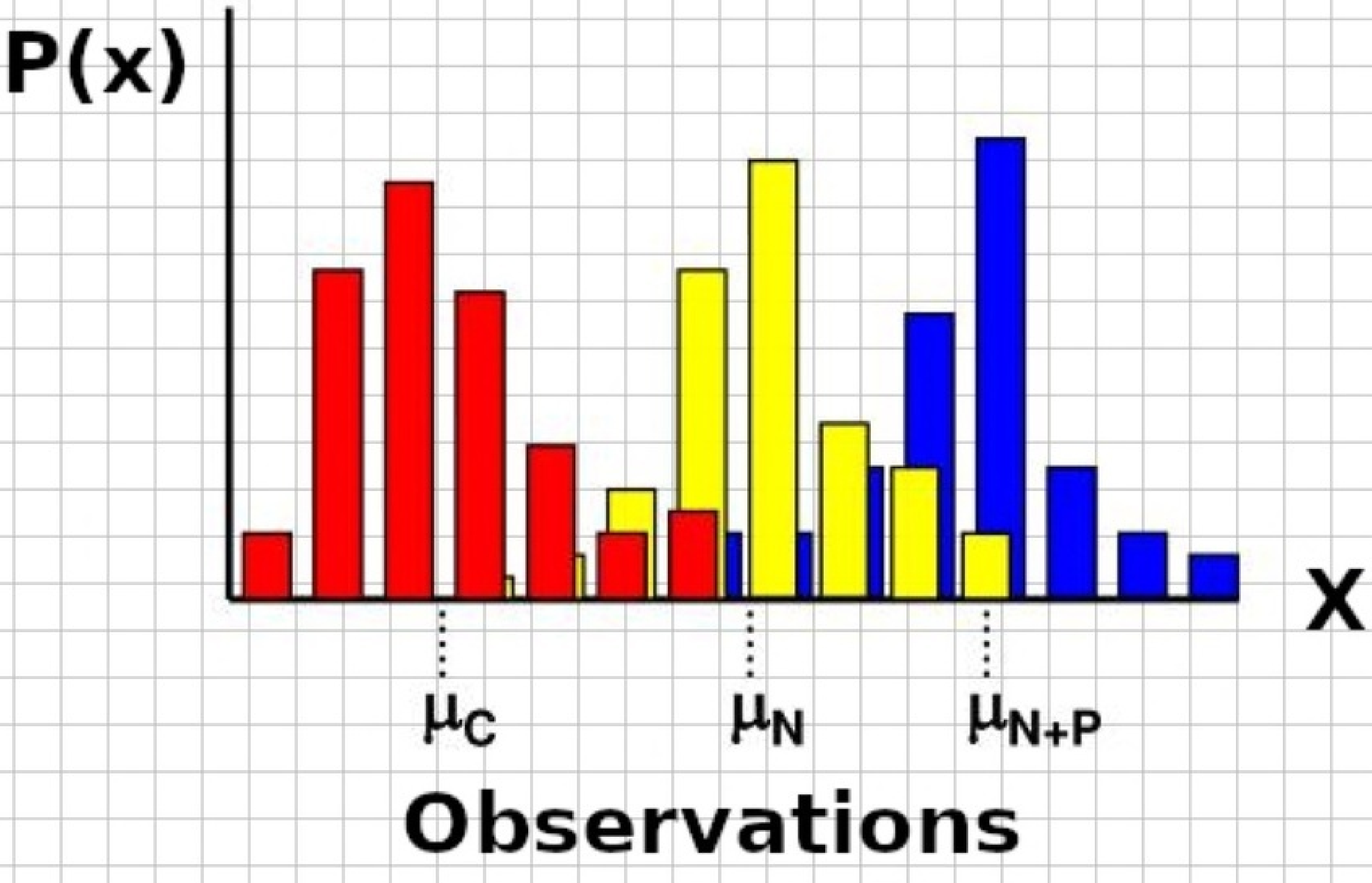
PrivateTeacher

Cours Privés de Science

Analyse de Variance ANOVA



Introduction

Lorsque l'on souhaite comparer la taille de deux individus, il est facile de faire la différence.

On mesure la taille du premier, puis celle du second, si elles sont identiques, les deux personnes ont la même taille, sinon, l'une des deux personnes est plus grande que l'autre.

Graphiquement, on a la situation suivante :

On interprète ce graphique facilement :

$$\text{taille}_{\text{Gimli}} < \text{taille}_{\text{Legolas}}$$

$\Rightarrow$ Gimli est plus petit que Legolas.

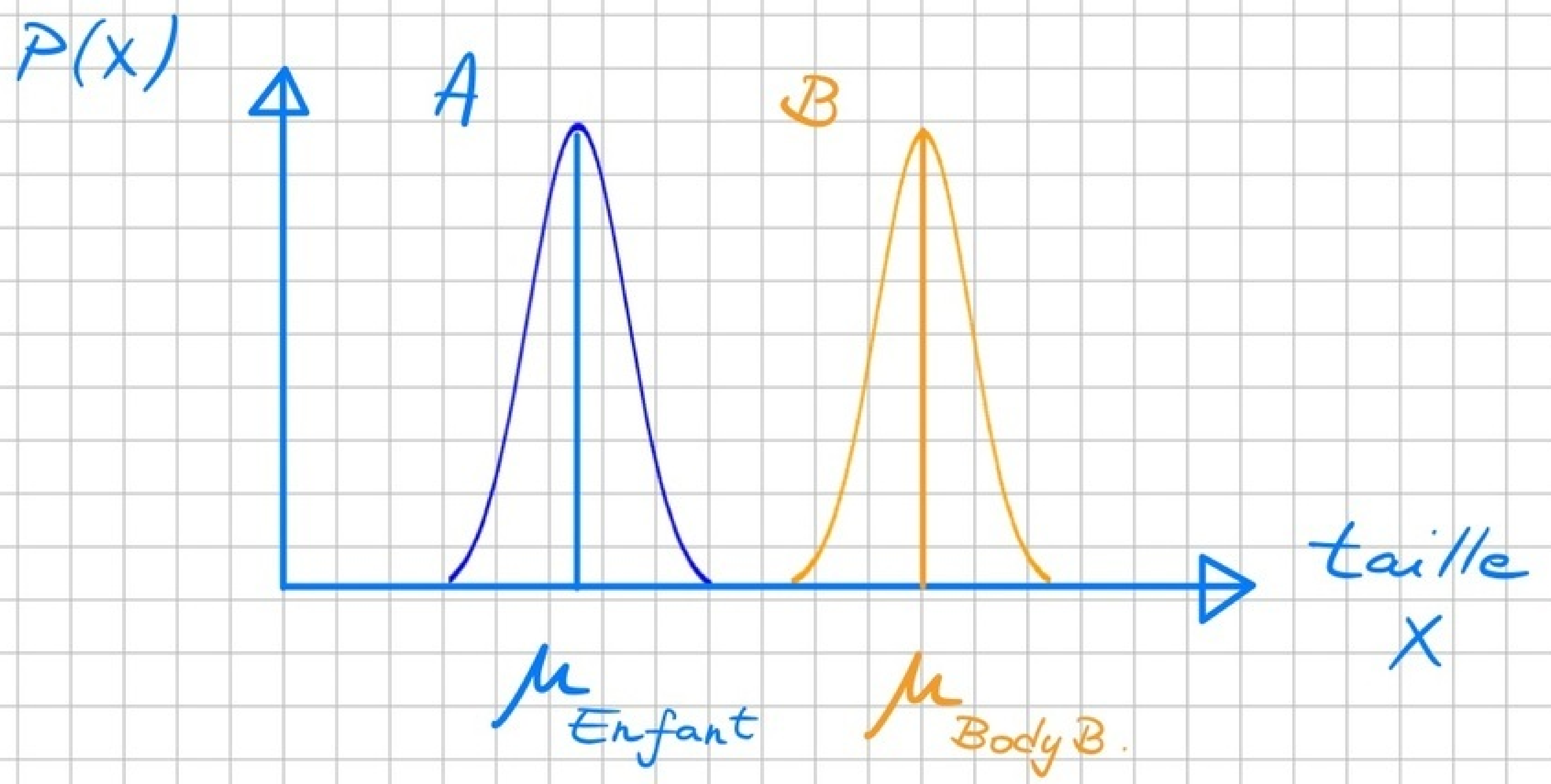
La réponse est simple, évidente, immédiate.

Admettons maintenant que l'on souhaite comparer non pas deux personnes, mais deux groupes.

Le gpe A : constitué d'enfants entre 5 et 6 ans.

Le gpe B : constitué de bodybuilders agé de 45 ans.

Graphiquement, on a la situation suivante :



Cette fois-ci on ne compare plus des tailles individuelles, mais deux ensembles de valeurs.

Le gpe A en effet contient des enfants plus petits et d'autres plus grands formant ainsi une variété de valeurs qui se distribuent de part et d'autre de la moyenne.

Une situation similaire se rencontre chez les bodybuilders.

Pour comparer la taille des deux groupes, on peut bien évidemment se servire de leur moyennes respectives.

On dira alors $\mu_A < \mu_B$

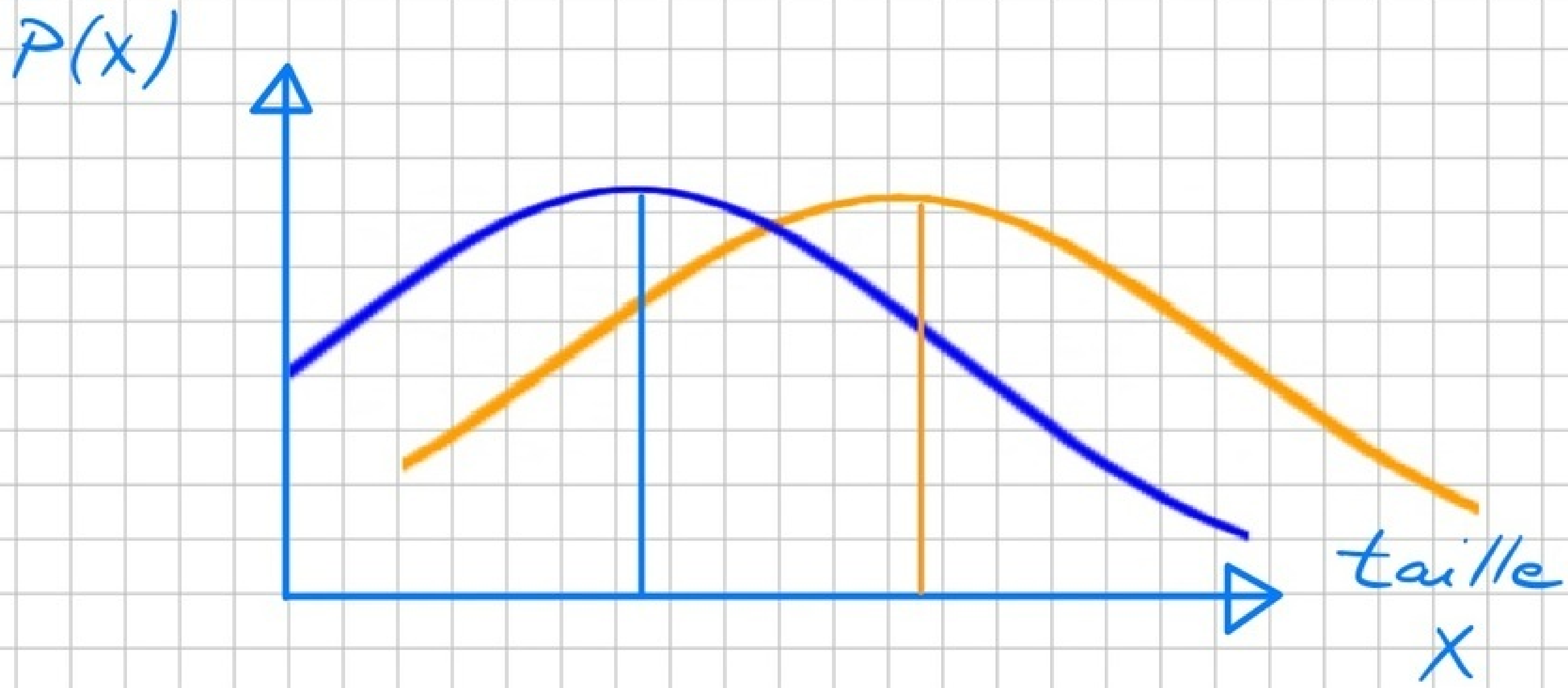
On peut également tenire le raisonnement suivant : Il n'existe aucun recouvrement entre les valeurs du gpe d'enfants avec les valeurs du gpe de bodybuilders.
Les deux gpes sont donc différents

Notez bien ceci cependant : Il est facile de faire la différence entre ces deux gpe car il s'agit de deux gpes bien distincts.

Que se passerait-il si la composition des gpes était plus variée ? Imaginons par exemple deux groupes constitués d'adultes aussi bien que d'enfants. Le premier gpe provient du canton de Fribourg et le deuxième du Valais.

Que se passerait-il si l'on tentait de comparer leur taille ?

Dans ce cas, nous aurions la situation suivante



Que peut-on dire à présent au sujet de la taille des groupes. Sont-elles vraiment différentes ? Peut-on se contenter de comparer leur moyenne respective ?

Leur moyenne sont certes différentes, mais le recouvrement des valeurs est important. La plupart des valeurs du premier groupe en effet se trouvent également dans le second, et réciproquement.

Se pourrait-il que cette différence de moyenne que l'on observe ici ne soit due qu'au hasard et que si l'on répétait les mesures la situation serait inversée ?

Les statisticiens confrontés à cette situation pour la première fois ont imaginés une solution ingénieuse.

Il s'agit de la solution que tu vas découvrir dans ce document. Elle repose sur l'analyse de la variance.

La variance elle même étant une <u>mesure de la dispersion</u> nous allons commencer par là.

Mesure de dispersion

On utilise la dispersion pour mesurer la diversité des observations au sein d'un groupe.

Il existe de nombeuses manières de mesurer la dispertion. On utilise le plus souvent la variance et l'écart-type.

Il en existe d'autres comme l'étendue ou la distance interquartile. Toutes sans exceptions donnent une mesure de la diversité des observations.

Lorsque la diversité des observations est grande, la variance est importante. On dit que le groupe est hétérogène.

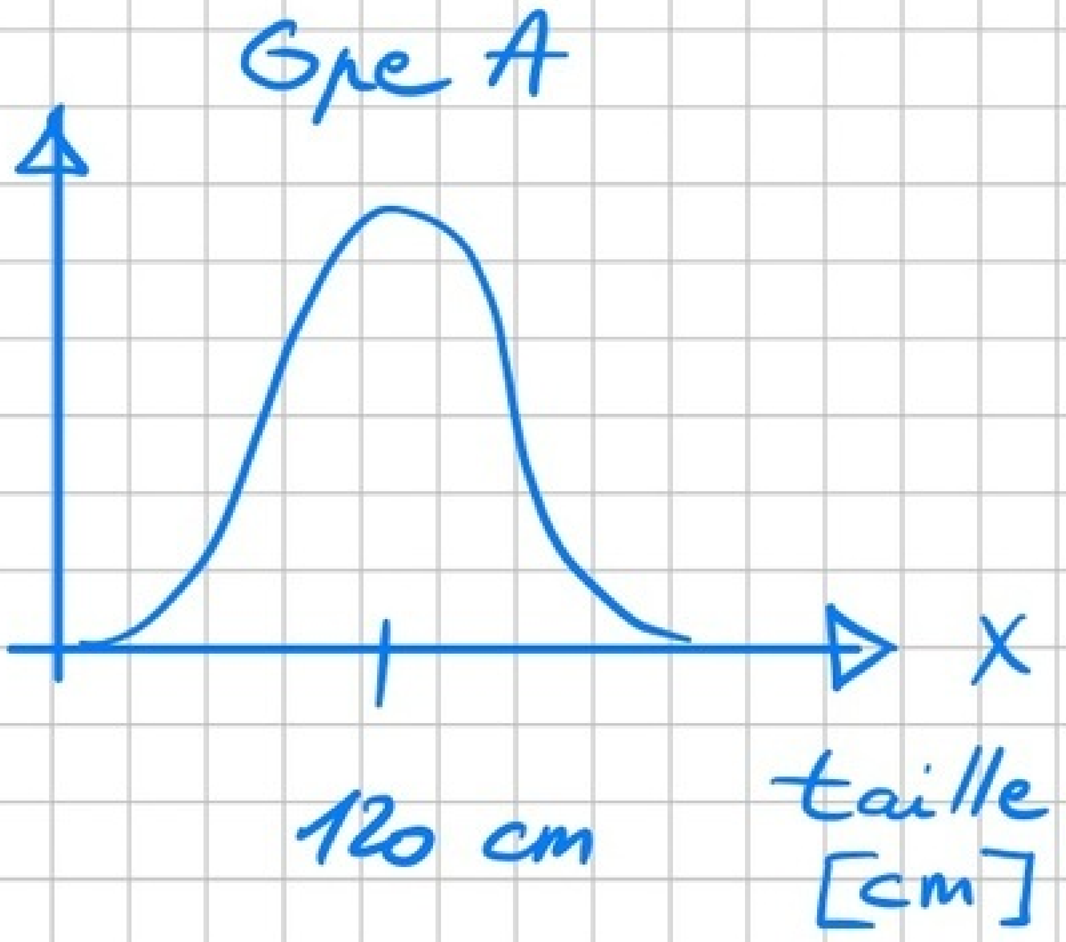
À l'inverse lorsque la diversité des observations est petite, la variance est faible. On dira alors que le gpe est homogène.

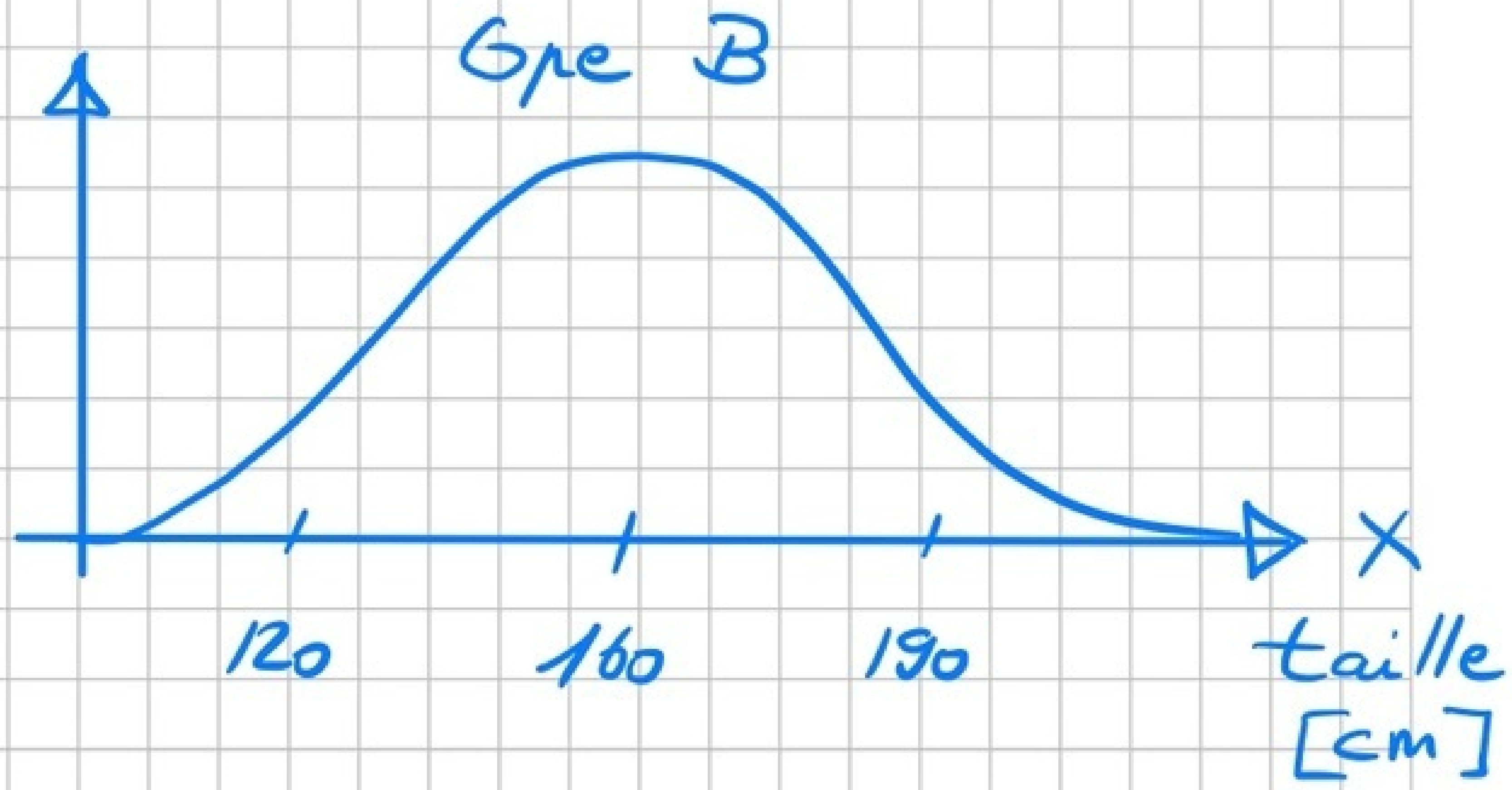
Imaginons deux groupes A et B.

Le groupe A est constitué par des enfant entre 5 et 6 ans.

Le groupe B est composé d'enfants d'adolescents et de jeunes adultes.

Si l'on s'intéresse maintenant à la taille des individus dans chacun de ces groupes, nous ferons les observations suivantes :





petite dispertion	grande dispertion
les tailles des individus sont semblables	les tailles des individus sont différentes les une des autres
le groupe est homogène	Le groupe est hétérogène.

$\boxed{\text{La dispertion est une mesure de la diversité des observations}}$

La variance

Intéressons-nous à présent à cette mesure de la dispertion que l'on nomme variance.

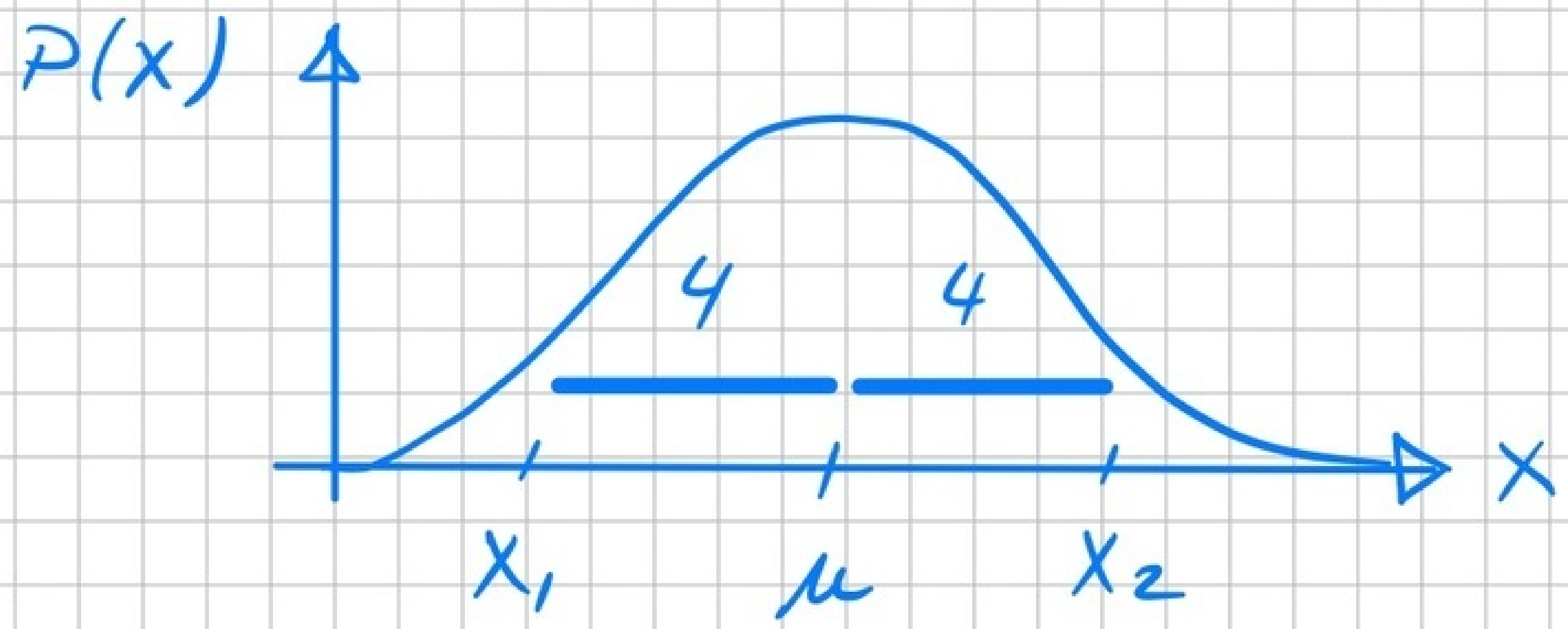
On la désigne par le signe σ^2
la lettre greque sigma : σ

La signification de cette grandeur se trouve dans son expression mathématique.

$$\boxed{Var(X) = \frac{1}{dl} \sum_i (x_i - \mu)^2}$$

Intéressons nous pour commencer à la deuxième partie de la formule : la somme des carrés.

Pour simplifier, intéressons-nous à seulement deux valeurs de X : X_1 à gauche de la moyenne et X_2 à droite.



On le voit, la variance mesure la somme des écarts entre les valeurs et leur moyenne

Les écarts sont mis au carré pour éviter que les écarts à droite de la moyenne ne compense les écarts à gauche.

Les écarts se compensent

$$(X_1 - \mu) + (X_2 - \mu) = (-4) + 4 = 0$$

les écarts s'accumulent

$$(X_1 - \mu)^2 + (X_2 - \mu)^2 = (-4)^2 + 4^2 = 32$$

De cette manière on à donc une mesure de la somme des écarts de l'ensemble des valeur par rapport à la moyenne du groupe.

Le tout est ensuite normalisé par le degrès de liberté $\frac{1}{dl}$ de l'échantillon.

$$Var(X) = \frac{1}{dl} \sum_i (X_i - \mu)^2$$

Au sujet des degrès de liberté

Le degrès de liberté est un concepte important en physique, en science en générale et en statistique en particulier. IL convient donc de s'y arrêter un instant.

Le degrès de liberté est une mesure de l'étendue des possibilités dont on dispose pour expliquer une observation.

Imaginez par exemple que vous êtes en colocation avec 3 personnes. Vous venez d'acheter un poulet que vous mettez au frigo puis vous partez dans votre chambre pour étudier.

À votre retour, le poulet a disparu :-(
Combiens de possibilités avez-vous pour expliquer cette observation ?

Vous l'avez compris, vous en avez trois. IL s'agit du nombre de personnes avec qui vous partagez l'apartement.

Le degrès de liberté est donc égale à 3

On note $dl = 3$

Imaginez maintenant que vous partagiez votre chambre avec une personne. Combiens de possibilités avez vous à présent pour expliquer la disparition du poulet ?

$$3 - 1 = 2$$

Le degrés de libertés diminue d'autant que vous avez de contraintes qui lie vos colocataire. Ici l'un d'eux fait partie du même groupe que vous. Vous savez donc tout de suite si c'est lui ou pas. L'étendue des possibilités à diminué le degrés de liberté est donc plus petit.

Différence entre deux moyennes

Nous avons vu que lorsque l'on souhaite comparer deux groupes, comparer leur moyenne ne suffit pas, encore faut-il tenir compte de leur variance.

Deux groupes ont des moyennes différentes $\mu_1 \neq \mu_2$

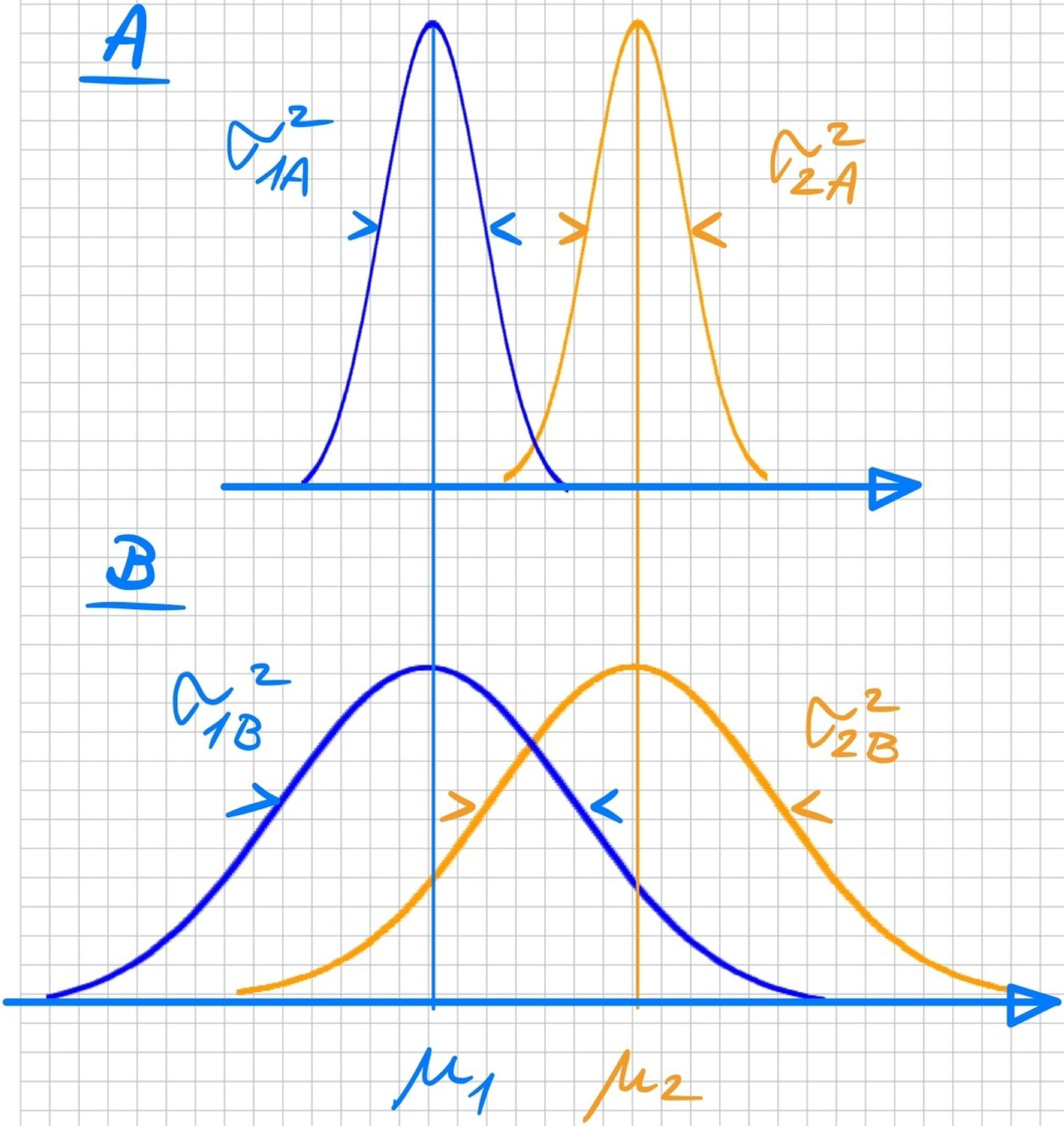
Imaginons deux groupes dont les moyennes sont différent imaginons deux scénarios différents :

Scénario A : la variance des deux groupes sont faibles.

Scénario B : la variance des deux groupes sont importantes

Dans les deux cas, la différence des moyennes est identique.

Voici une représentation de la situation :



Effet de la variance

Dans le cas A, aucune des valeurs du premier groupe ne se retrouve dans le deuxième. Il n'y a aucun recouvrement il est donc claire que les deux groupes sont différents.

Dans le cas B, le recouvrement est important. Il devient difficile de distinguer les valeurs du premier groupe avec celles du second. La différence entre les deux groupes n'est plus si évidente.

En fait, la différence entre les groupes s'évanoui à mesure que la variance augmente.

Pour mesurer la différence entre les moyennes, on a donc imaginé la statistique suivante :

$$\boxed{t = \frac{\mu_1 - \mu_2}{f(\sigma_1^2 ; \sigma_2^2)}}$$

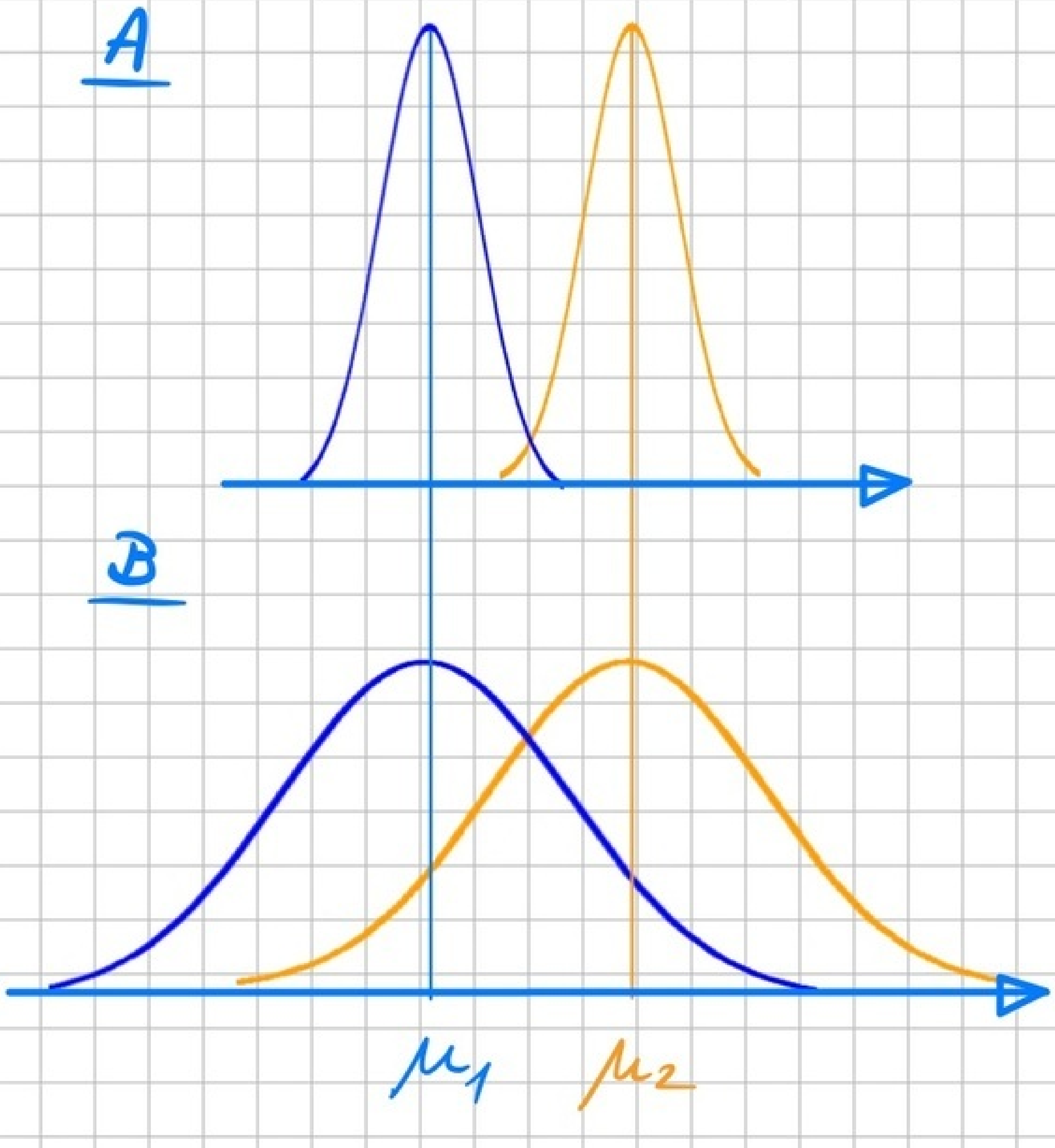
$f(\sigma_1^2 ; \sigma_2^2)$ est une fonction de la variance σ^2 des groupes 1 et 2

La valeur de $f(\sigma_1^2 ; \sigma_2^2)$ est d'autant plus grande que la variance des groupes 1 et 2 est importante.

σ_1^2 et/ou σ_2^2 grand

$\Rightarrow f(\sigma_1^2 ; \sigma_2^2)$ grand

> Ce qui fait la différence entre deux populations, ce n'est pas la différence entre leur moyenne mais <u>le rapport</u> de cette différence avec leur variance.



$\mu_1 - \mu_2$ grand <u>par rapport à</u> σ_1^2 et σ_2^2

$\mu_1 - \mu_2$ petit <u>par rapport à</u> σ_1^2 et σ_2^2

Plutôt que d'utiliser simplement la différence entre les moyennes, on utilise le rapport entre cette différence et la variance.

t est une mesure de ce rapport

t petit signifie que les groupes sont identiques

t grand signifie que les groupes sont différent

Entre les deux, on fix une valeur limite qui nous permet de faire la différence

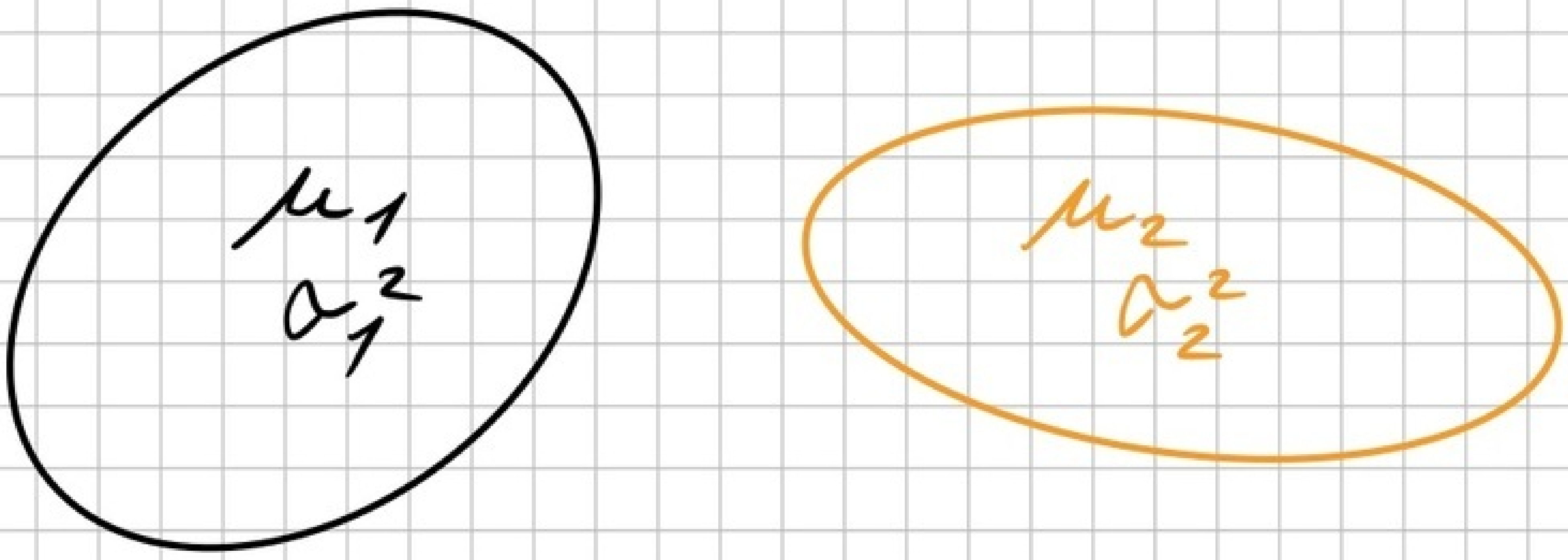
On appel cette valeur limite de t : t_{crit}

t est donc petit s'il est plus petit que t_{crit}
t est grand s'il est plus grand que t_{crit}

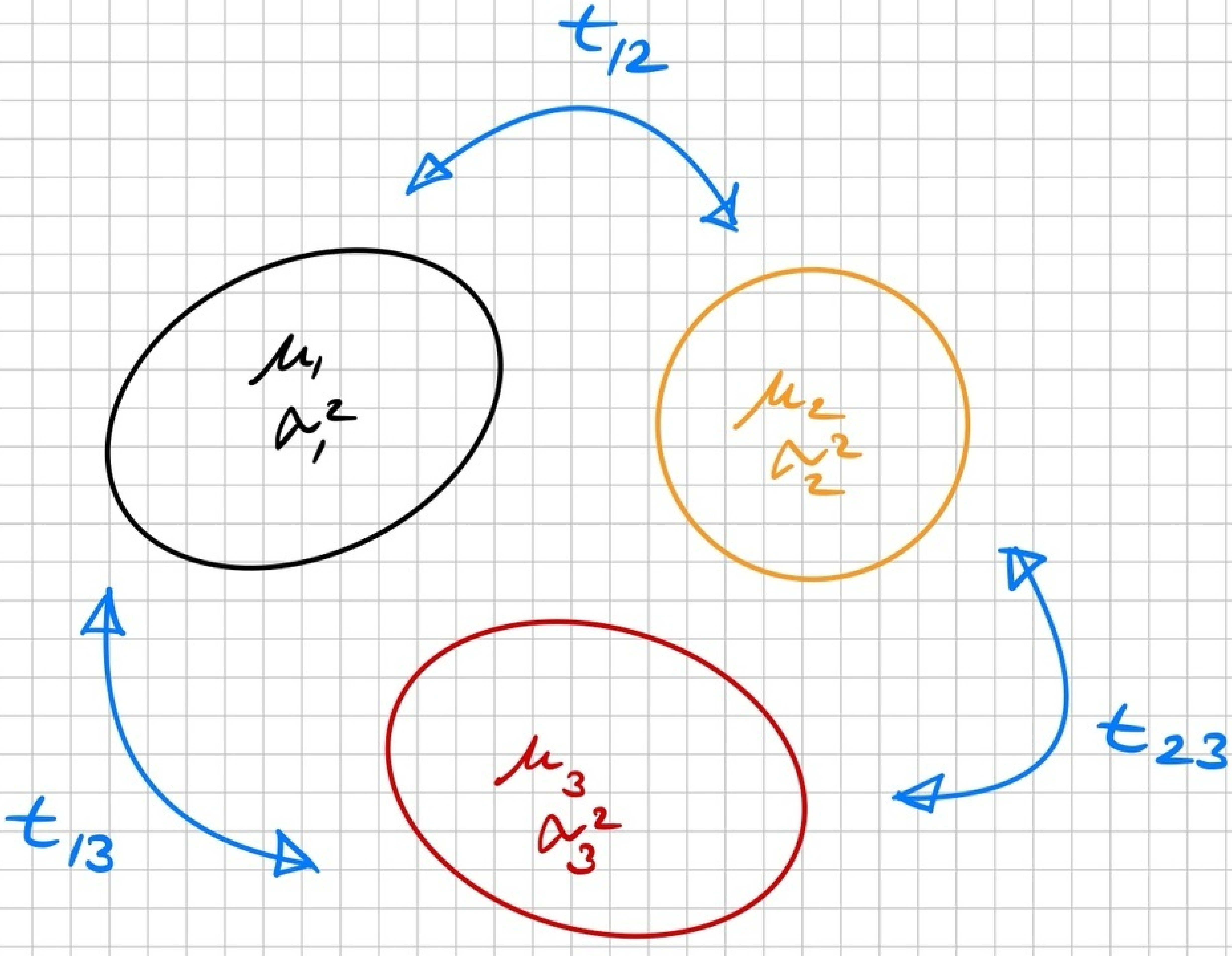
$t < t_{crit}$
$\Rightarrow$ gpe identique

$t > t_{crit}$
$\Rightarrow$ gpes différent

Lorsqu'on souhaite comparer les moyennes de deux groupes, on peut donc utiliser t comme statistique de test et réaliser un test de Student :



La même démarche peut ainsi être répétée autant de fois que nécessaire pour tester tous les groupe entre eux



Pourquoi une ANOVA

Cette démarche cependante est fastidieuse lorsque le nombre de groupes est important

Dans ce cas, les statisticiens. nes ont imaginé une procédure qui compare tous les groupes en une fois.

ANOVA repose sur une statistique qui ressemble a la mesure t, mais qui regroupe les variance de tous les groupes à la fois.

On appel cette mesure la statistique F

De même que nous l'avons vu pour la variance, sa signification repose dans son expression mathématique :

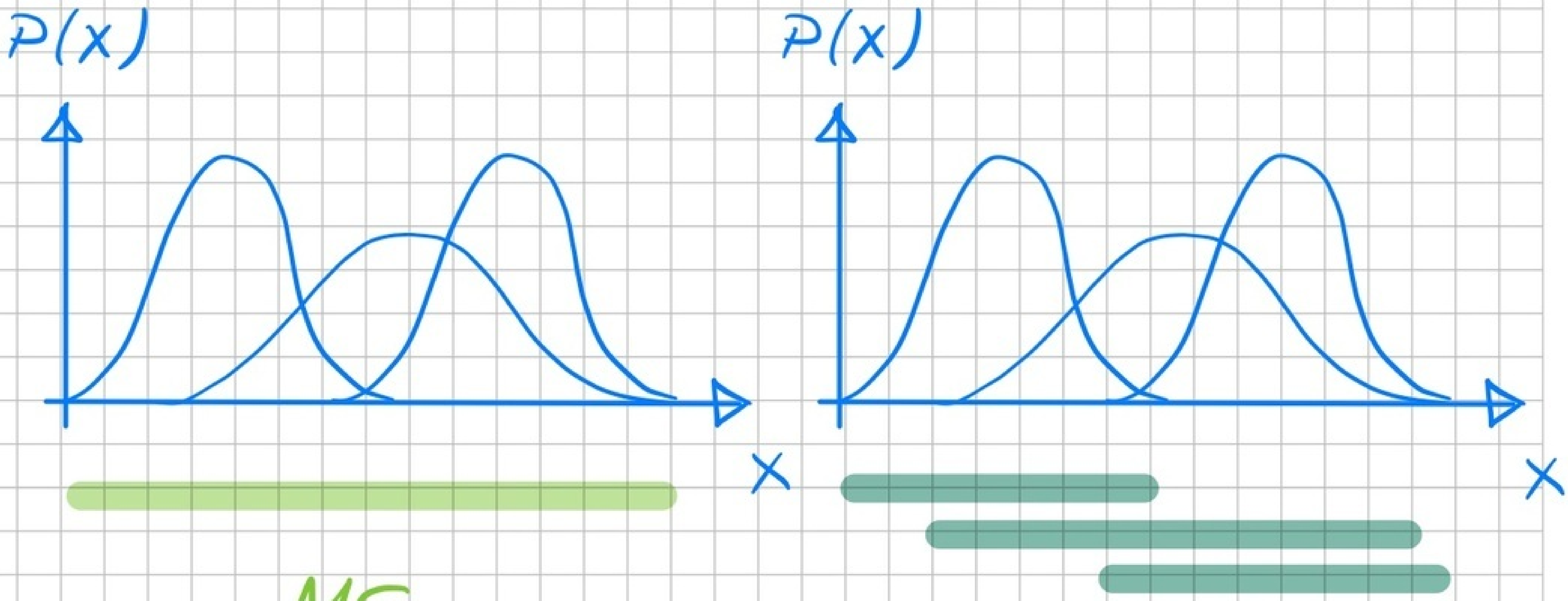
$$\boxed{F = \frac{MC_{inter}}{MC_{intra}}}$$

Comme pour t, il s'agit d'un rapport entre deux valeurs

MC_{inter} représente la moyenne des carrés <u>entre</u> les gpes

MC_{intra} représente la moyenne des carré <u>au sein</u> des gpes

Imaginons 3 groupes nommés A, B et C dont les moyennes valent μ_A, μ_B et μ_C respectivement. Graphiquement, on a la représentation suivante :



MC_{inter}

Dispertion de l'ensemble des groupes

MC_{intra}

Moyenne de la dispertion de chaque groupe

$$\boxed{F = \frac{MC_{inter}}{MC_{intra}}}$$

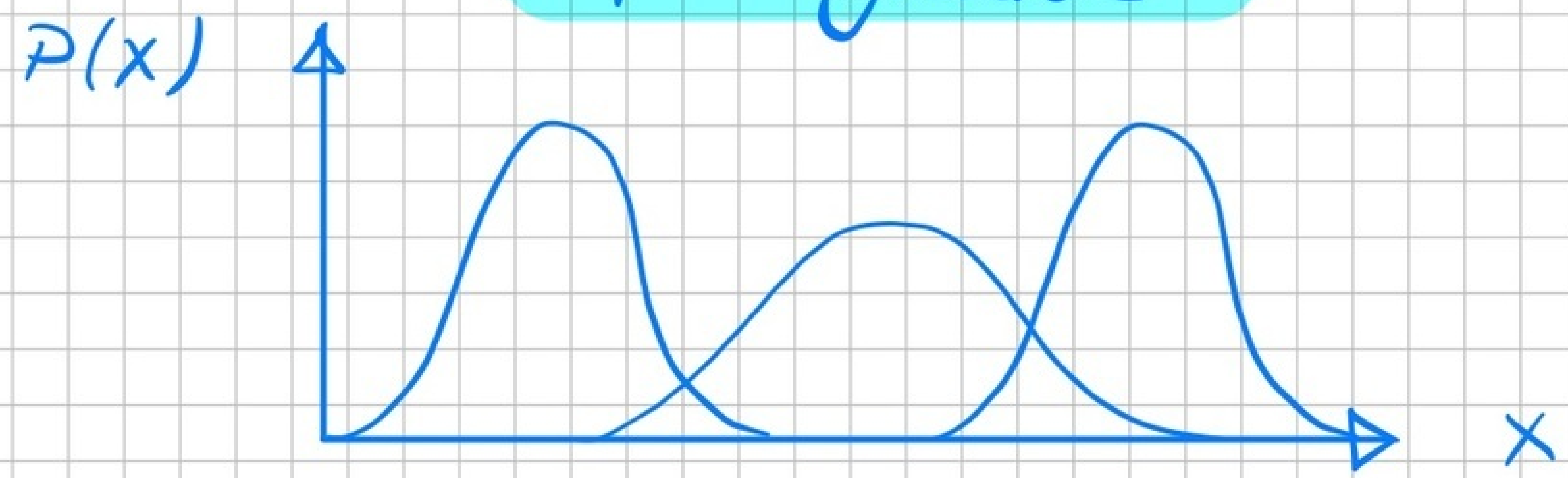
Cette valeur F est le rapport entre la dispertion entre les groupes MC_{inter} et la dispertion au sein des groupe MC_{intra}

Deux scénario sont alors possible :

1) L'espacement entre les groupes MC_{inter} est important par rapport à la dispertion au sein des groupes MC_{intra}

$$\boxed{MC_{inter} > MC_{intra}}$$

F grand

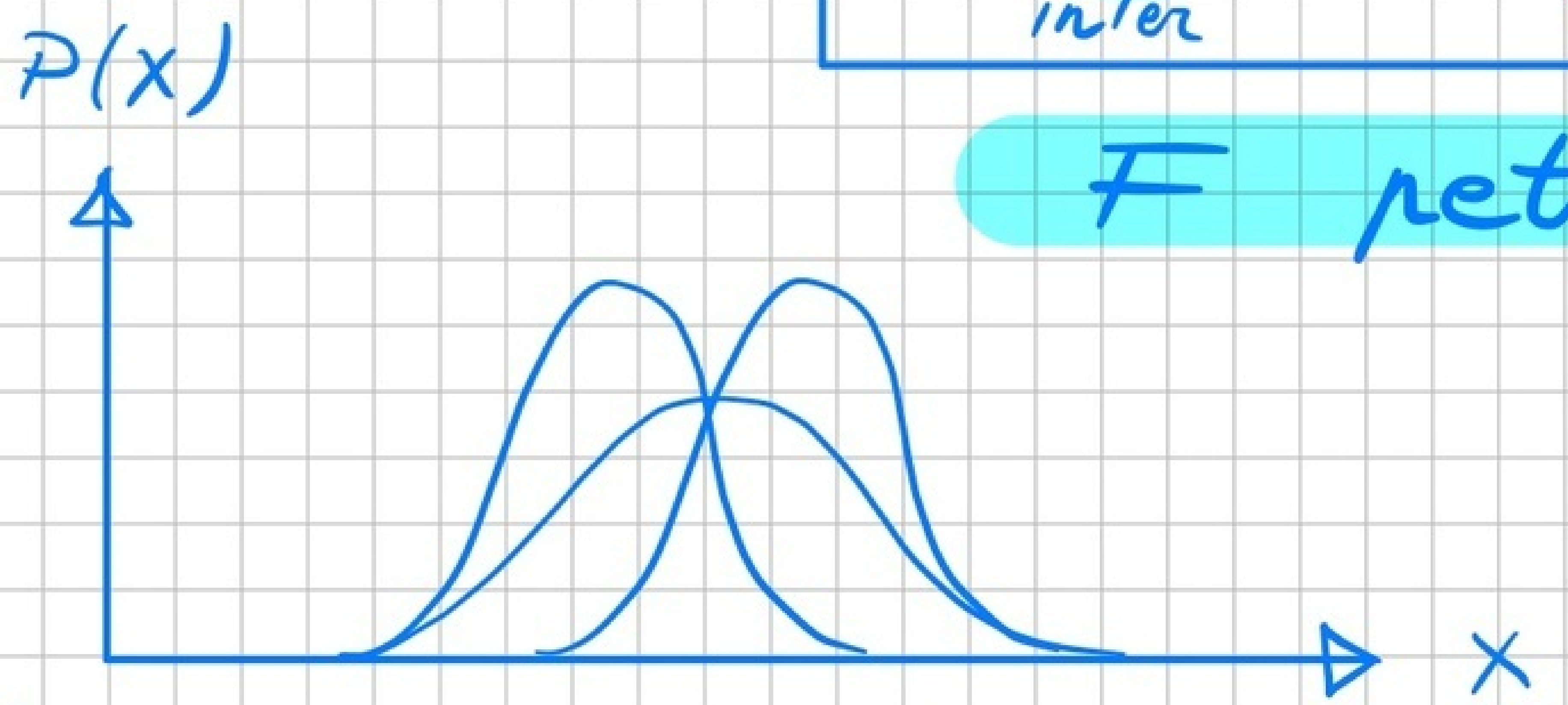


Les groupes sont différents

2) L'espacement entre les groupes MC_{inter} est petit par rapport à la dispertion au sein des groupes MC_{intra}

$$\boxed{MC_{inter} < MC_{intra}}$$

F petit



Les groupes sont semblables

Pour calculer F, on calculera donc chacune de ces deux valeurs, MC_{inter} et MC_{intra}, l'une après l'autre.

Voici la démarche :

1) Pour calculer F, on commence par calculer la moyenne des observations dans chaque groupe, puis la moyenne des moyennes que l'on nomme " grande moyenne " μ_G

2) On calcul ensuite les écarts.

   a) les écarts entre chaque observation et leur moyenne donne SC_{intra}

   b) les écarts entre chaque moyenne et la grande moyenne donne SC_{inter}

3) On calcul ensuite la moyenne des écarts

   a) $MC_{inter} = \dfrac{SC_{inter}}{dl_{inter}}$

   b) $MC_{intra} = \dfrac{SC_{intra}}{dl_{intra}}$

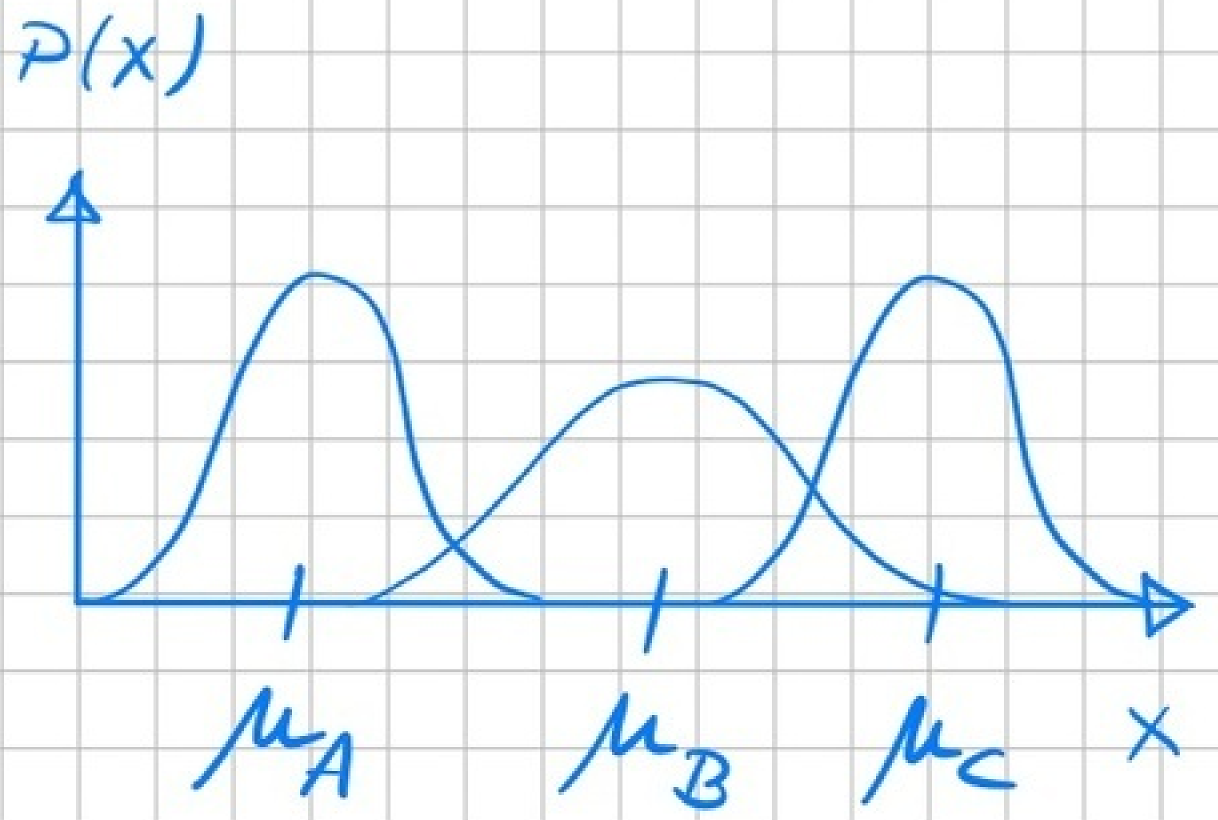
4) Le rapport entre ces moyennes finalement donne la statistique F.

$$F = \frac{MC_{inter}}{MC_{intra}}$$

MC_{intra}

Dispertion au sein d'un groupe

Intra - Groupe



1a) Calculer les moyennes de chaque gpe

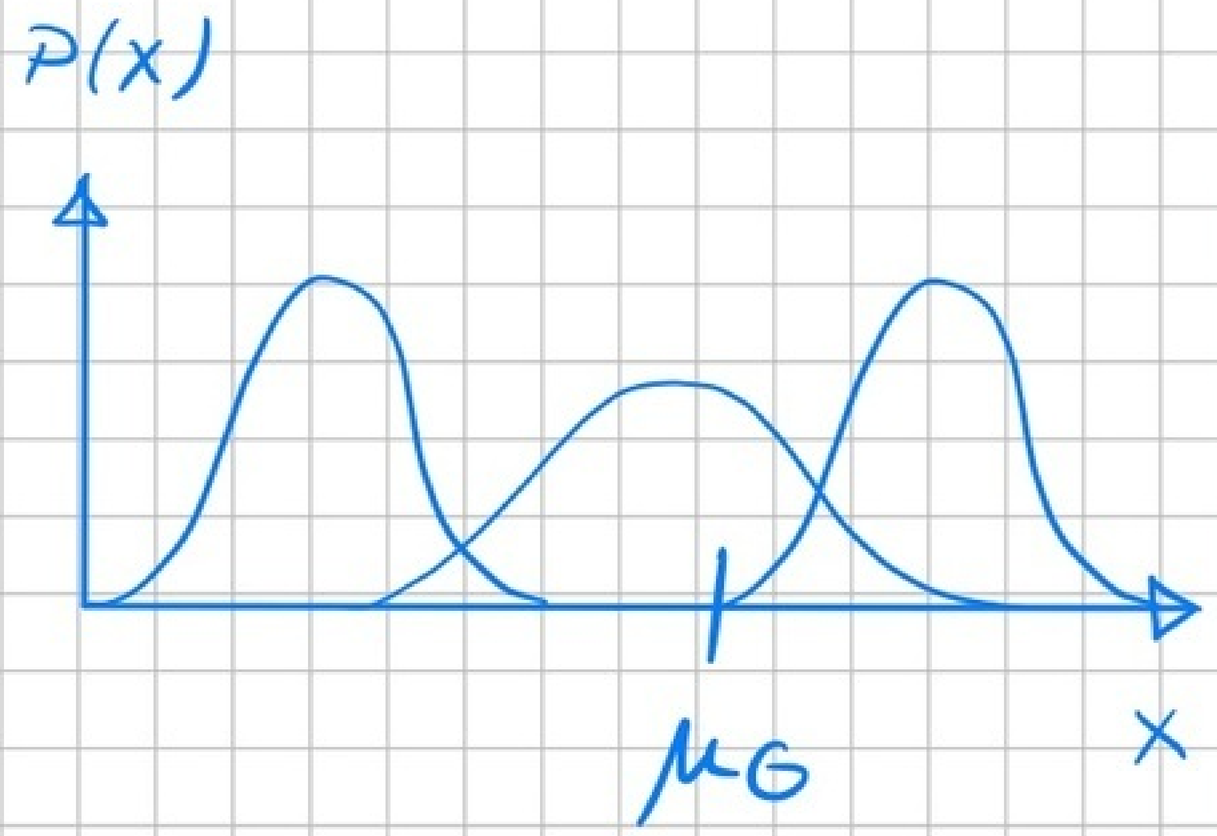
2a) Calculer la somme des écart SC_{intra}

3a) Calculer la moyenne des ecarts MC_{intra}

MC_{inter}

Dispertion entre les groupes

Inter - groupe



1b) Calculer la moyenne générale de l'ensemble des groupes

2b) Calculer la somme des écart SC_{inter}

3b) Calculer la moyenne des ecarts MC_{inter}

4) $F = \dfrac{MC_{inter}}{MC_{intra}}$

Un exemple de calcul

Imaginons des mesures faite sur trois groupes différents. On reporte nos observation dans le tableau suivant :

Gpe A	Gpe B	Gpe C
5	5	4
5	4	3
4	4	3
4	3	2
3	3	1

Suivant la démarche énoncée précédemment, on commence par calculer la moyenne de chaque groupe :

1a) $\mu_A = \frac{1}{5}(5+5+4+4+3) = 4.2$

$\mu_B = \frac{1}{5}(5+4+4+3+3) = 3.8$

$\mu_C = \frac{1}{5}(4+3+2+2+1) = 2.4$

2a) On calcul ensuite les écarts entre chaque observations et leur moyenne :

Gpe A	Gpe B	Gpe C
$5 - \mu_A$	$5 - \mu_B$	$4 - \mu_C$
$5 - \mu_A$	$4 - \mu_B$	$3 - \mu_C$
$4 - \mu_A$	$4 - \mu_B$	$3 - \mu_C$
$4 - \mu_A$	$3 - \mu_B$	$2 - \mu_C$
$3 - \mu_A$	$3 - \mu_B$	$1 - \mu_C$

Le carré des écarts nous donne la somme des carrés SC de chaque groupe.

Gpe A	Gpe B	Gpe C
$(5 - 4.2)^2$	$(5 - 3.8)^2$	$(4 - 2.4)^2$
$(5 - 4.2)^2$	$(4 - 3.8)^2$	$(3 - 2.4)^2$
$(4 - 4.2)^2$	$(4 - 3.8)^2$	$(3 - 2.4)^2$
$(4 - 4.2)^2$	$(3 - 3.8)^2$	$(2 - 2.4)^2$
$(3 - 4.2)^2$	$(3 - 3.8)^2$	$(1 - 2.4)^2$
SC_A	SC_B	SC_C

$$SC_A = (5-4.2)^2 + (5-4.2)^2 + (4-4.2)^2 + (4-4.2)^2 + (3-4.2)^2$$

$$= 0.64 + 0.64 + 0.4 + 0.4 + 1.44$$

$$= \underline{2.8}$$

$$SC_B = (5-3.8)^2 + (4-3.8)^2 + (4-3.8)^2 + (3-3.8)^2 + (3-38)^2$$

$$= 1.44 + 0.4 + 0.4 + 0.64 + 0.64$$

$$= \underline{2.8}$$

$$SC_C = (4-2.4)^2 + (3-2.4)^2 + (3-2.4)^2 + (2-2.4)^2 + (1-2.4)^2$$

$$= 2.56 + 0.36 + 0.16 + 0.16 + 1.96$$

$$= \underline{5.2}$$

La somme des carrés inter-groupe est donné par

$$\underline{SC_{intra}} = SC_A + SC_B + SC_C$$

$$= 2.8 + 2.8 + 5.2$$

$$= \underline{\mathbf{10.8}}$$

3a) On calcul enfin la moyenne des carré intra-groupe :

$$MC_{intra} = \frac{SC_{intra}}{dl_{intra}}$$

dl_{intra} est le degrès de liberté intra-groupe

$$dl_{intra} = N - k$$

N = nombre d'observation
k = nombre de groupe

$$dl_{intra} = 15 - 3 = 12$$

$$MC_{intra} = \frac{SC_{intra}}{dl_{intra}}$$

$$= \frac{10.8}{12}$$

$$= 0.9$$

$$MC_{intra} = 0.9$$

1b) Pour calculer la grande moyenne μ_G on calcul la moyenne des moyennes

$$\mu_G = \frac{1}{3}\left(\mu_A + \mu_B + \mu_C\right)$$

$$\mu_A = 4.2$$
$$\mu_B = 3.8$$
$$\mu_C = 2.4$$

$$= \frac{1}{3}\, 10.4 = 3.47$$

$$\underline{\mu_G = 3.47}$$

2b) le calcul de MC_{inter} repose sur la somme des écarts entre les moyennes de chaque groupe et la grande moyenne :

Pour chaque groupe on a

Ecart gpe A : $E_A = (\mu_A - \mu_G)^2$

$$= (4.2 - 3.47)^2$$

$$\underline{= 0.53}$$

Ecart gpe B : $E_B = (\mu_B - \mu_G)^2$

$$= (3.8 - 3.47)^2$$

$$\underline{= 0.11}$$

Ecart gpe C : $E_C = (\mu_C - \mu_G)^2$

$$= (2.4 - 3.47)^2$$

$$\underline{= 1.47}$$

La somme des carrés inter-groupe est la somme des écarts de chaque groupe pondérée par le nombre d'observation "n" de chaque groupe

$$n_A = 5$$
$$n_B = 5$$
$$n_C = 5$$

$$\underline{SC_{inter}} = n_A E_A + n_B E_B + n_C E_C$$

$$= 5 \cdot 0.53 + 5 \cdot 0.11 + 5 \cdot 1.14$$

$$= 2.65 + 0.55 + 5.7$$

$$\underline{= 8.9}$$

3b) La moyenne des carrés inter-groupe enfin se calcul ainsi :

$$MC_{inter} = \frac{SC_{inter}}{dl_{inter}}$$

dl_{inter} est le degrès de liberté inter-groupe

$$dl_{inter} = k - 1$$

k = nombre de groupe

$$dl_{inter} = 3 - 1 = 2$$

$$MC_{inter} = \frac{SC_{inter}}{dl_{inter}}$$

$$= \frac{8.9}{2}$$

$$= 4.45$$

$$MC_{inter} = 4.45$$

4) Finalement nous disposons de toutes les valeurs nécessaires pour calculer F

$$F = \frac{MC_{inter}}{MC_{intra}}$$

$$MC_{inter} = 4.45$$

$$MC_{intra} = 0.9$$

$$F = \frac{4.45}{0.9} = \underline{4.94}$$

$$F = 4.94$$

Tableau ANOVA

On regroupe ensuite tous ces résultats dans un tableau :

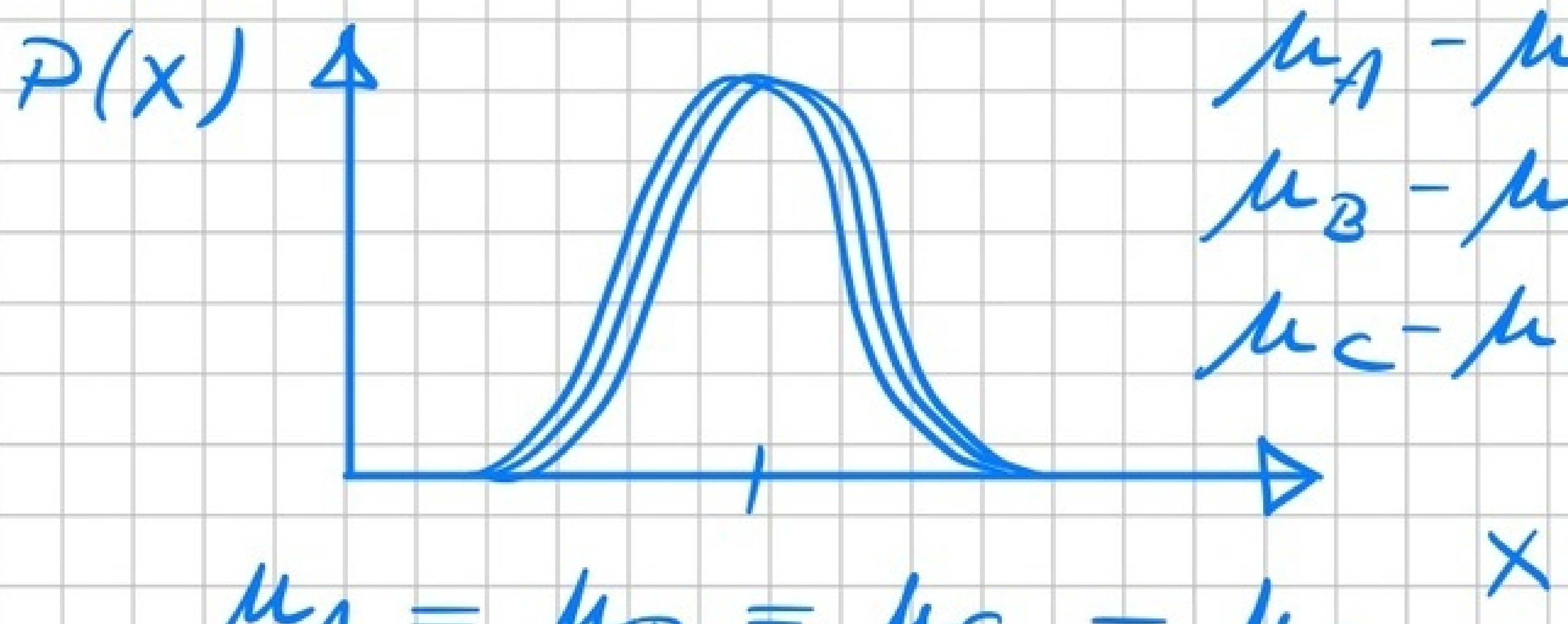
Variation	SC	dl	CM
Inter-Groupe	SC_{inter}	dl_{inter}	CM_{inter}
Intra-Groupe	SC_{intra}	dl_{intra}	CM_{intra}
Totale	SC_{tot}	dl_{tot}	CM_{tot}

Pour notre exemple, on à donc le tableau suivant :

Variation	SC	dl	MC
Inter-Groupe	8.9	2	4.45
Intra-Groupe	10.8	12	0.9
Totale	19.7	14	5.35

Interpréter la valeur F

Dans le cas où tous les groupes sont confondu, les valeurs SC_{inter} vaut 0



$$\mu_A - \mu_G = 0$$
$$\mu_B - \mu_G = 0$$
$$\mu_C - \mu_G = 0$$

$$\mu_A = \mu_B = \mu_C = \mu_G$$

- $SC_{inter} = \sum_{k}^{ABC} n_k (\mu_k - \mu_G)^2 = 0$
- $MC_{inter} = \dfrac{SC_{inter}}{dl_{inter}} = 0$
- $F = \dfrac{MC_{inter}}{MC_{intra}} = 0$

Pour des groupes identique, la valeur de F est donc proche de zéro.

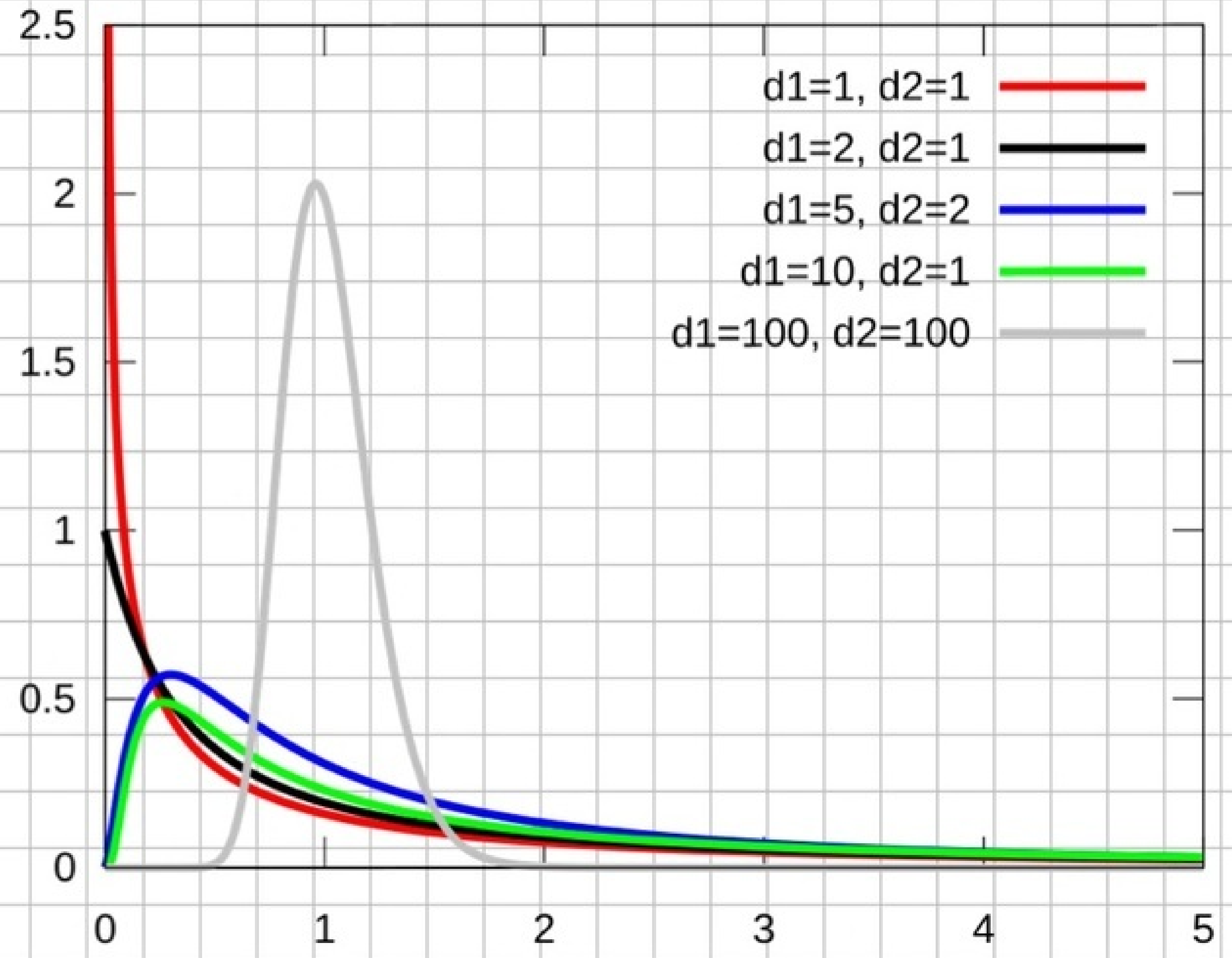
On peut donc se servir de F comme d'une mesure de la distance entre les groupes

Quelles sont les valeurs suffisamment loin de zéro pour considérer les groupes différent ?

Seul la distribution des valeurs de F peut nous le dire.

Distribution des valeurs de F

Les valeurs de F se distribuent selon une loi de Fisher. La forme de la distribution est donnée par les deux degrés de liberté dl_{inter} et dl_{intra}



Source: Wikipedia

dans notre exemple :

$$DF1 = dl_{inter} = 2$$
$$DF2 = dl_{intra} = 12$$

Ces 5% sont notre seuil α. Il détermine la valeur critique de F (F_{crit}) au dela de laquelle les valeurs de F sont considérées extrêmes et donc les groupes diffèrent !

Elles sont données dans les tables

Table des quantiles de la distribution F

Valeurs critique de F pour les distributions de Fisher au seuil $\alpha = 5\%$

	DF1	$\alpha = 0.05$						
DF2	1	2	3	4	5	6	7	8
1	161.45	199.5	215.71	224.58	230.16	233.99	236.77	238.88
2	18.513	19	19.164	19.247	19.296	19.33	19.353	19.371
3	10.128	9.5521	9.2766	9.1172	9.0135	8.9406	8.8867	8.8452
4	7.7086	6.9443	6.5914	6.3882	6.2561	6.1631	6.0942	6.041
5	6.6079	5.7861	5.4095	5.1922	5.0503	4.9503	4.8759	4.8183
6	5.9874	5.1433	4.7571	4.5337	4.3874	4.2839	4.2067	4.1468
7	5.5914	4.7374	4.3468	4.1203	3.9715	3.866	3.787	3.7257
8	5.3177	4.459	4.0662	3.8379	3.6875	3.5806	3.5005	3.4381
9	5.1174	4.2565	3.8625	3.6331	3.4817	3.3738	3.2927	3.2296
10	4.9646	4.1028	3.7083	3.478	3.3258	3.2172	3.1355	3.0717
11	4.8443	3.9823	3.5874	3.3567	3.2039	3.0946	3.0123	2.948
12	4.7472	3.8853	3.4903	3.2592	3.1059	2.9961	2.9134	2.8486
13	4.6672	3.8056	3.4105	3.1791	3.0254	2.9153	2.8321	2.7669
14	4.6001	3.7389	3.3439	3.1122	2.9582	2.8477	2.7642	2.6987
15	4.5431	3.6823	3.2874	3.0556	2.9013	2.7905	2.7066	2.6408
16	4.494	3.6337	3.2389	3.0069	2.8524	2.7413	2.6572	2.5911
17	4.4513	3.5915	3.1968	2.9647	2.81	2.6987	2.6143	2.548
18	4.4139	3.5546	3.1599	2.9277	2.7729	2.6613	2.5767	2.5102
19	4.3807	3.5219	3.1274	2.8951	2.7401	2.6283	2.5435	2.4768
20	4.3512	3.4928	3.0984	2.8661	2.7109	2.599	2.514	2.4471
21	4.3248	3.4668	3.0725	2.8401	2.6848	2.5727	2.4876	2.4205
22	4.3009	3.4434	3.0491	2.8167	2.6613	2.5491	2.4638	2.3965
23	4.2793	3.4221	3.028	2.7955	2.64	2.5277	2.4422	2.3748
24	4.2597	3.4028	3.0088	2.7763	2.6207	2.5082	2.4226	2.3551
25	4.2417	3.3852	2.9912	2.7587	2.603	2.4904	2.4047	2.3371
26	4.2252	3.369	2.9752	2.7426	2.5868	2.4741	2.3883	2.3205
27	4.21	3.3541	2.9604	2.7278	2.5719	2.4591	2.3732	2.3053
28	4.196	3.3404	2.9467	2.7141	2.5581	2.4453	2.3593	2.2913
29	4.183	3.3277	2.934	2.7014	2.5454	2.4324	2.3463	2.2783
30	4.1709	3.3158	2.9223	2.6896	2.5336	2.4205	2.3343	2.2662
40	4.0847	3.2317	2.8387	2.606	2.4495	2.3359	2.249	2.1802
60	4.0012	3.1504	2.7581	2.5252	2.3683	2.2541	2.1665	2.097
120	3.9201	3.0718	2.6802	2.4472	2.2899	2.175	2.0868	2.0164
Inf	3.8415	2.9957	2.6049	2.3719	2.2141	2.0986	2.0096	1.9384

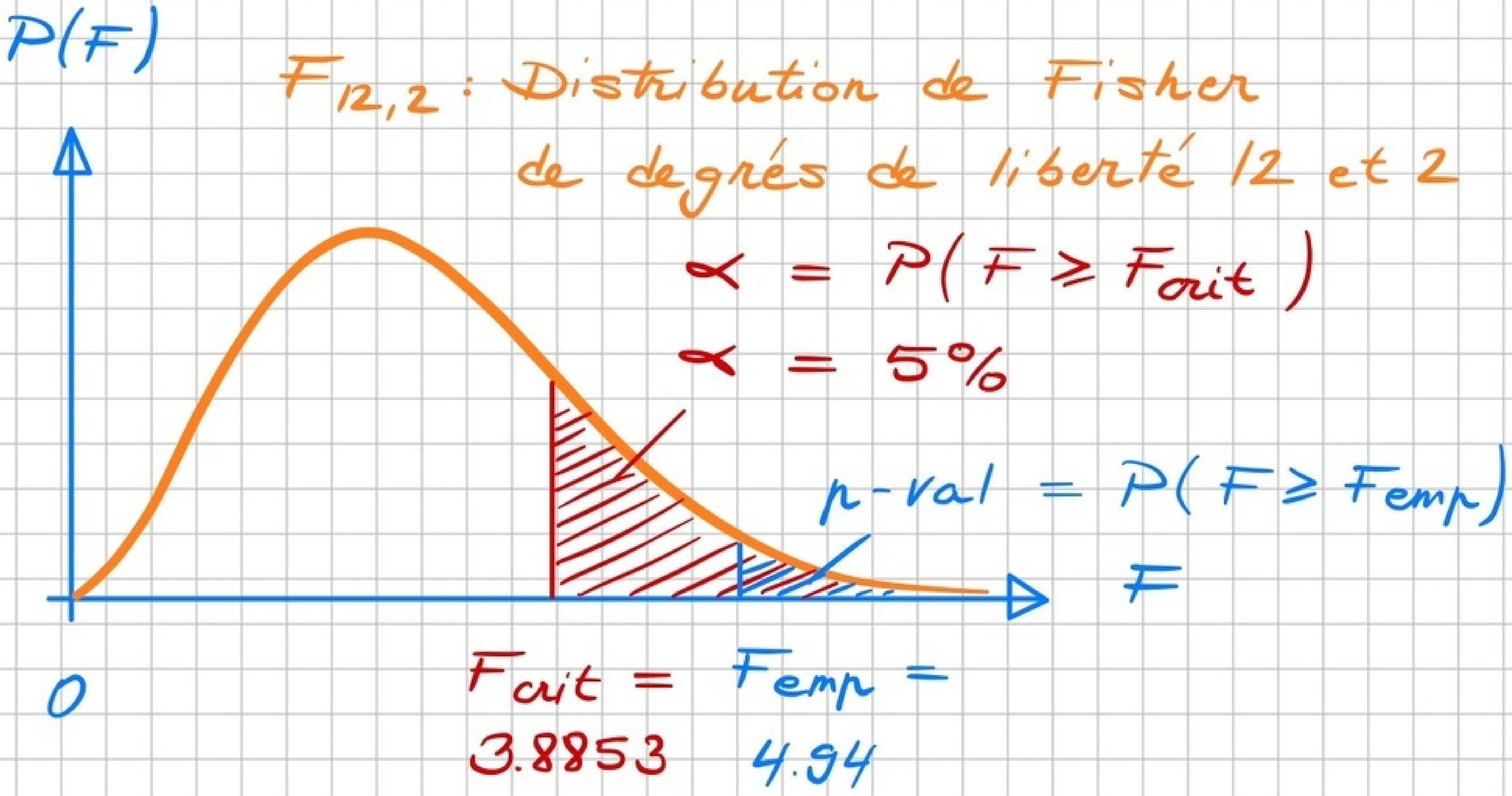
On lit dans la table que le F critique F_{crit} au seil 5% pour une distribution de Fisher dont les degrés de liberté valent 12 et 2 est de 3.8853

$$F_{crit} = 3.8853$$

La valeur de F calculée précédemment est une valeur issue de l'expérience. On la nomme F empirique F_{emp}

$$F_{emp} = 4.94$$

On a la représentation suivante



Interprétation

La valeur empirique de F, F_{emp} calculée à partir de nos observations est plus grande que la valeur critique de F F_{crit} définie par notre seuil α de 5%

Nous avons donc un élément en faveur de l'hypothèse selon laquelle les trois groupes auxquels appartiennent nos observations sont différents.

À ce stade, on ne peut pas encore dire quels groupes sont différents. Seul des tests post-hoc nous permettrons de savoir.

La statistique de test F cependant nous permet de dire qu'il existe une différence statistiquement significative.

C'est ainsi que l'on peut distinger les moyennes les unes des autres sur la base de la variance de chacun des groupes

C'est ce qui s'appel une analyse de variance ANOVA