FORMATIONS AUX STATISTIQUES : MANUEL DU FORMATEUR DE NIVEAU 1

Janvier 1991

SCÉNARIO 10

Utilisation d'un test statistique : comparaison de 2 échantillons

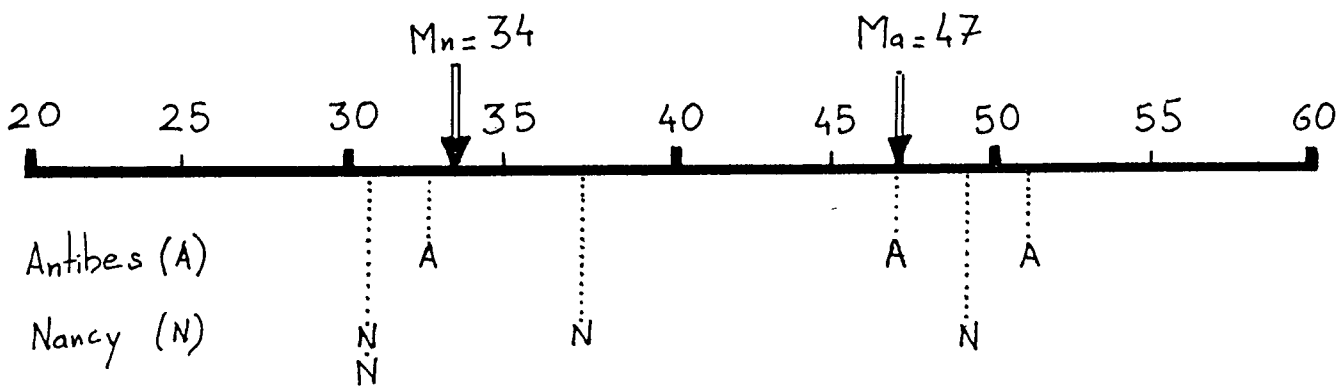
INRA

TRANSPARENTS

DÉPARTEMENT DE BIOMÉTRIE

formation permanente FP

AGES DES AGENTS DES CENTRES D'ANTIBES ET DE NANCY CALCUL DES MEDIANES ET DES SOMMES DES RANGS



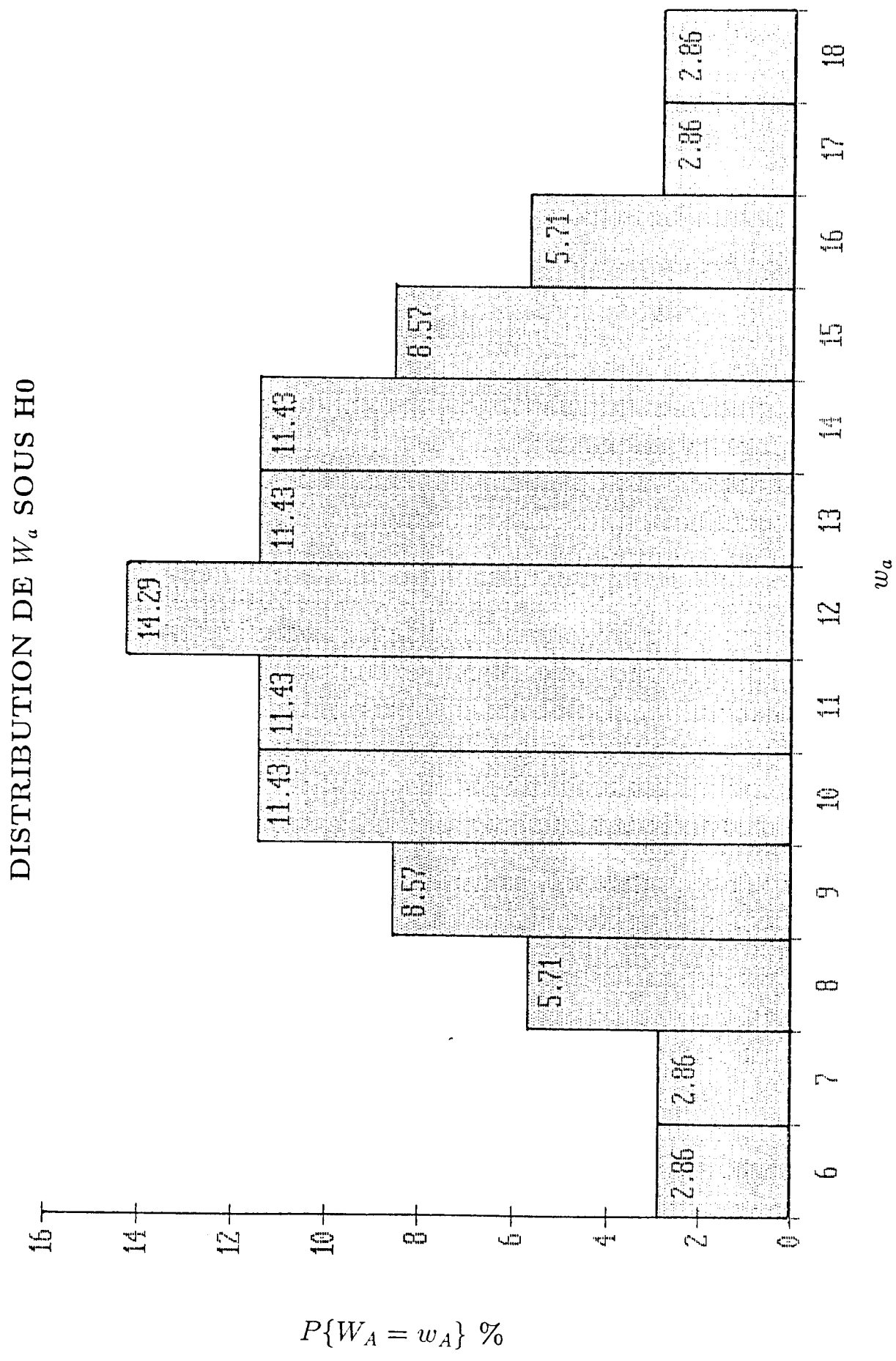
CENTRE	AGE DES AGENTS				MÉDIANE
Antibes	47	51	33	20	47
Nancy	37	49	31	31	34

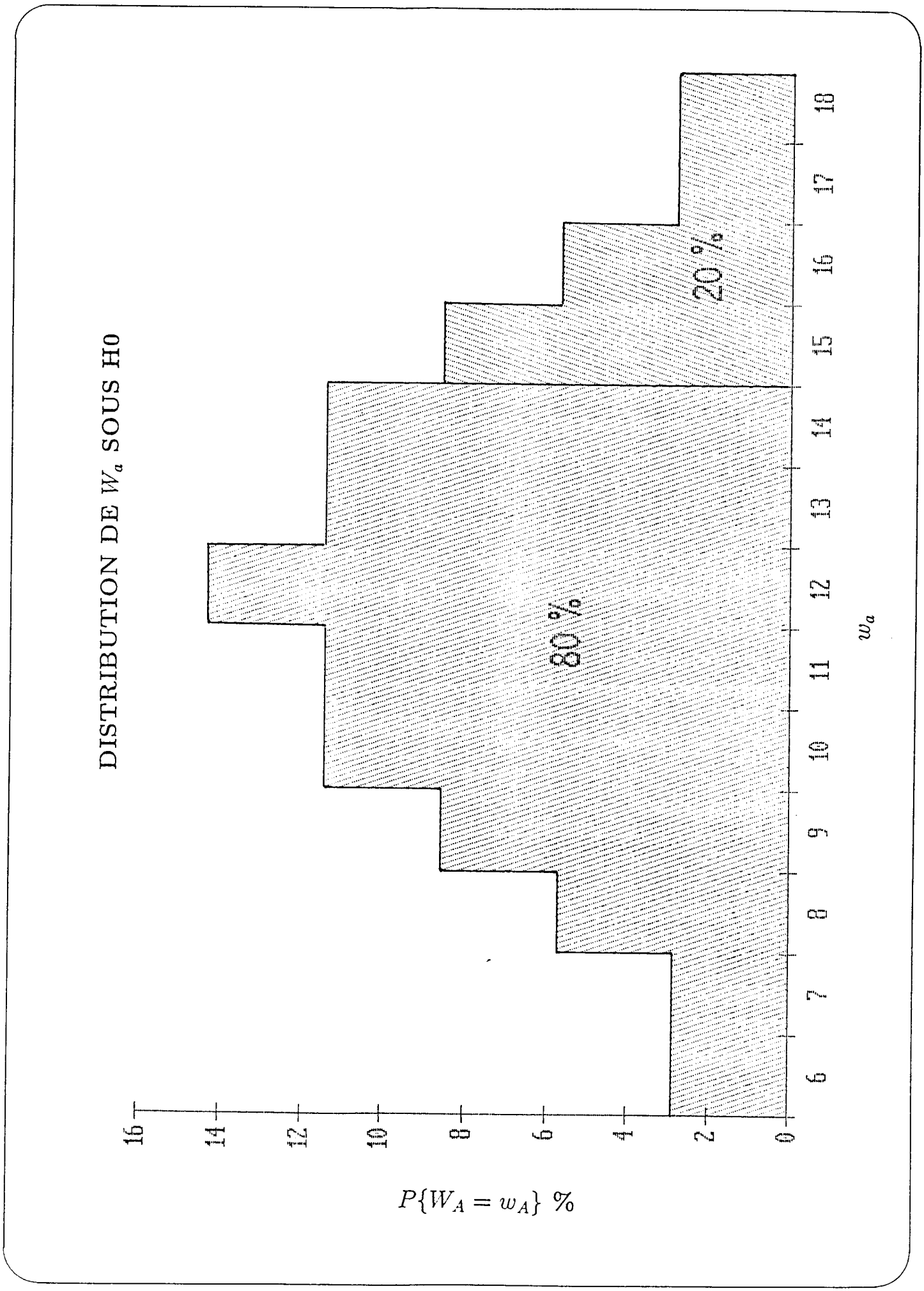
	RANGS DE A	RANGS DE N
	3	1,5
	5	1,5
	7	4
		6
Somme des rangs	Wa = 15	Wn = 13

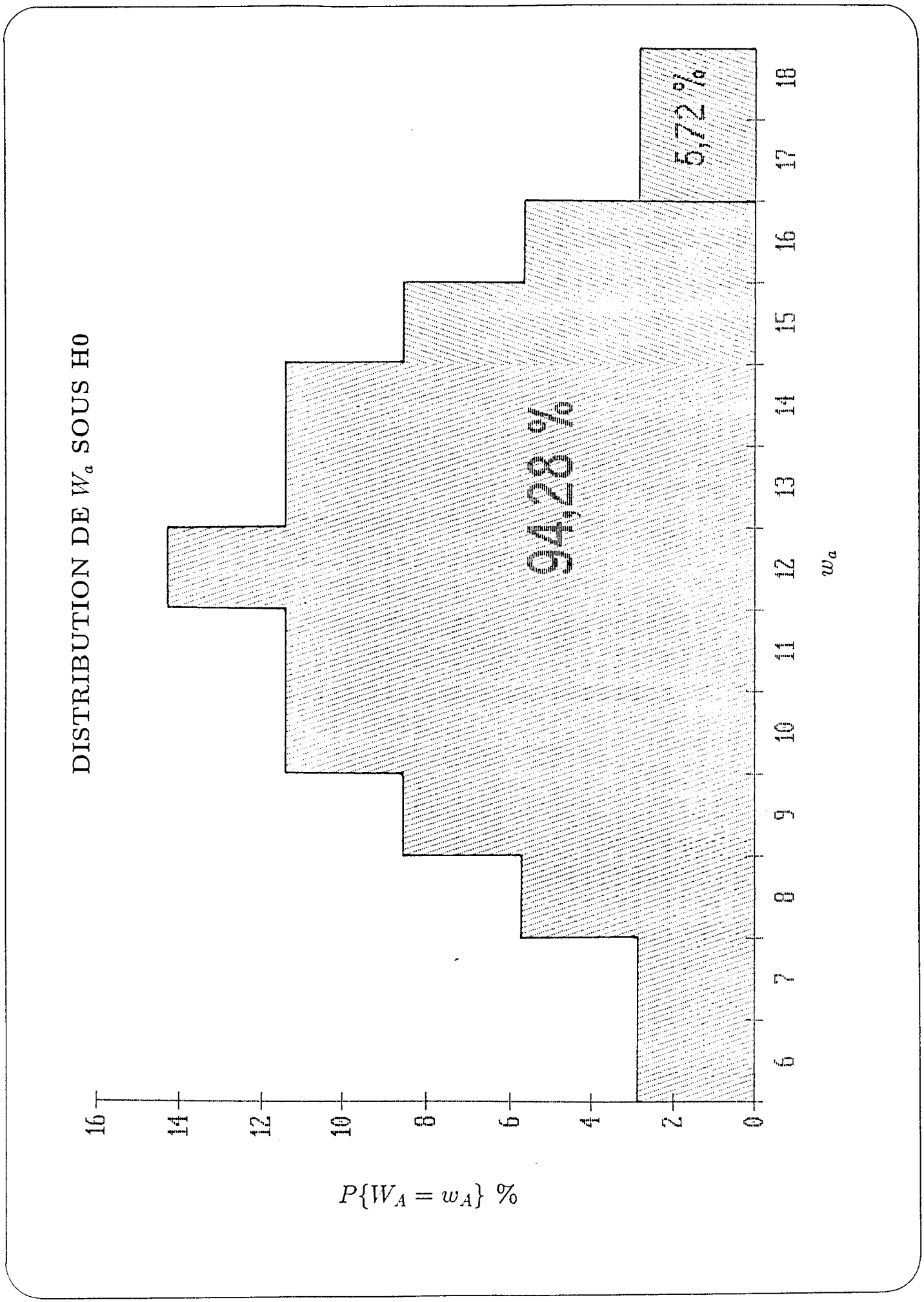
TABLEAU DES PROBABILITÉS CUMULÉES DE W_a SOUS $H0$

Liste des 35 combinaisons possibles — Rangs 1	2	3	4	5	6	7	Valeurs possibles des Sommes des Rangs wn	wa	Nombre de combinaisons donnant w_a	Loi de probabilité de W_a en % $P\{W_a = w_A\}$	Probabilité Cumulée de W_a en % $P\{W_a \geq w_a\}$
N	N	N	N	A	A	A	10	18	1/35	2.86%	2.86%
N	N	N	A	N	A	A	11	17	1/35	2.86%	5.71%
N	N	A	N	N	A	A	12	16			
N	N	N	A	A	N	A	12	16	2/35	5.71%	11.43%
N	A	N	N	N	A	A	13	15			
N	N	A	N	A	N	A	13	15			
N	N	N	A	A	A	N	13	15	3/35	8.57%	20.00%
A	N	N	N	N	A	A	14	14			
N	A	N	N	A	N	A	14	14			
N	N	A	A	N	N	A	14	14			
N	N	A	N	A	A	N	14	14	4/35	11.43%	31.43%
A	N	N	N	A	N	A	15	13			
N	A	N	A	N	N	A	15	13			
N	A	N	N	A	A	N	15	13			
N	N	A	A	N	A	N	15	13	4/35	11.43%	42.86%
A	N	N	A	N	N	A	16	12			
N	A	A	N	N	N	A	16	12			
A	N	N	N	A	A	N	16	12			
N	A	N	A	N	A	N	16	12			
N	N	A	A	A	N	N	16	12	5/35	14.29%	57.14%
A	N	A	N	N	N	A	17	11			
A	N	N	A	N	A	N	17	11			
N	A	A	N	N	A	N	17	11			
N	A	N	A	A	N	N	17	11	4/35	11.43%	68.57%
A	A	N	N	N	N	A	18	10			
A	N	A	N	N	A	N	18	10			
A	N	N	A	A	N	N	18	10			
N	A	A	N	A	N	N	18	10	4/35	11.43%	80.00%
A	A	N	N	N	A	N	19	9			
A	N	A	N	A	N	N	19	9			
N	A	A	A	N	N	N	19	9	3/35	8.57%	88.57%
A	A	N	N	A	N	N	20	8			
A	N	A	A	N	N	N	20	8	2/35	5.71%	94.29%
A	A	N	A	N	N	N	21	7	1/35	2.86%	97.14%
A	A	A	N	N	N	N	22	6	1/35	2.86%	100.00%
Nb de Combinaisons:						35					

W_a est la valeur observée, w_A une valeur possible, c'est-à-dire successivement 6, 7, 8 ... 18


DISTRIBUTION DE W_a SOUS H0
16
14
12
10
8
6
4
2
0
2,86
2,86
5,71
8,57
11,43
11,43
14,29
11,43
11,43
8,57
5,71
2,86
2,86
6
7
8
9
10
11
12
13
14
15
16
17
18
w_a
$P\{W_A = w_A\}$ %


DISTRIBUTION DE W_a SOUS H0
80 %
20 %
w_a
$P\{W_A = w_A\}$ %
0
2
4
6
8
10
12
14
16
6
7
8
9
10
11
12
13
14
15
16
17
18


DISTRIBUTION DE W_a SOUS H0
94,28 %
5,72 %
w_a
$P\{W_A = w_A\}$ %
0
2
4
6
8
10
12
14
16
6
7
8
9
10
11
12
13
14
15
16
17
18

	Ce que l'on décide	
Situation Réelle	ACCEPTER H_0	REJETER H_0 (ou accepter H_1)
H_0 Vraie	décision correcte $1-\alpha$ ou niveau de confiance	erreur de 1ère espèce α
H_0 Fausse ou (H_1 vraie)	erreur de 2ème espèce β	décision correcte $1-\beta$ ou puissance du test

Résultats possibles d'un test d'hypothèse

LES ÉTAPES D'UN TEST D'HYPOTHÈSE

1. Définition des hypothèses H_o et H_1
2. Choix d'un résumé des données
   (ici rangs et sommes des rangs)
3. Chercher toutes les valeurs possibles de la somme des rangs sous H_o et leur probabilité dans la pratique on trouve ces valeurs dans une table.
4. Choisir le risque de 1ère espèce α
5. Faire le calcul de la statistique (ici W_a) et décider le rejet ou l'acceptation de l'hypothèse de départ en fonction soit d'une valeur critique (W_c) ou d'une probabilité critique.

FORMULE DU t DE STUDENT

$$t = \frac{|\hat{m}_1 - \hat{m}_2|}{Sd}$$

avec

$$Sd = S\sqrt{\frac{1}{n_1} + \frac{1}{n_2}}$$

et

$$S = \sqrt{\frac{(n_1 - 1)S_1^2 + (n_2 - 1)S_2^2}{n_1 + n_2 - 2}}$$

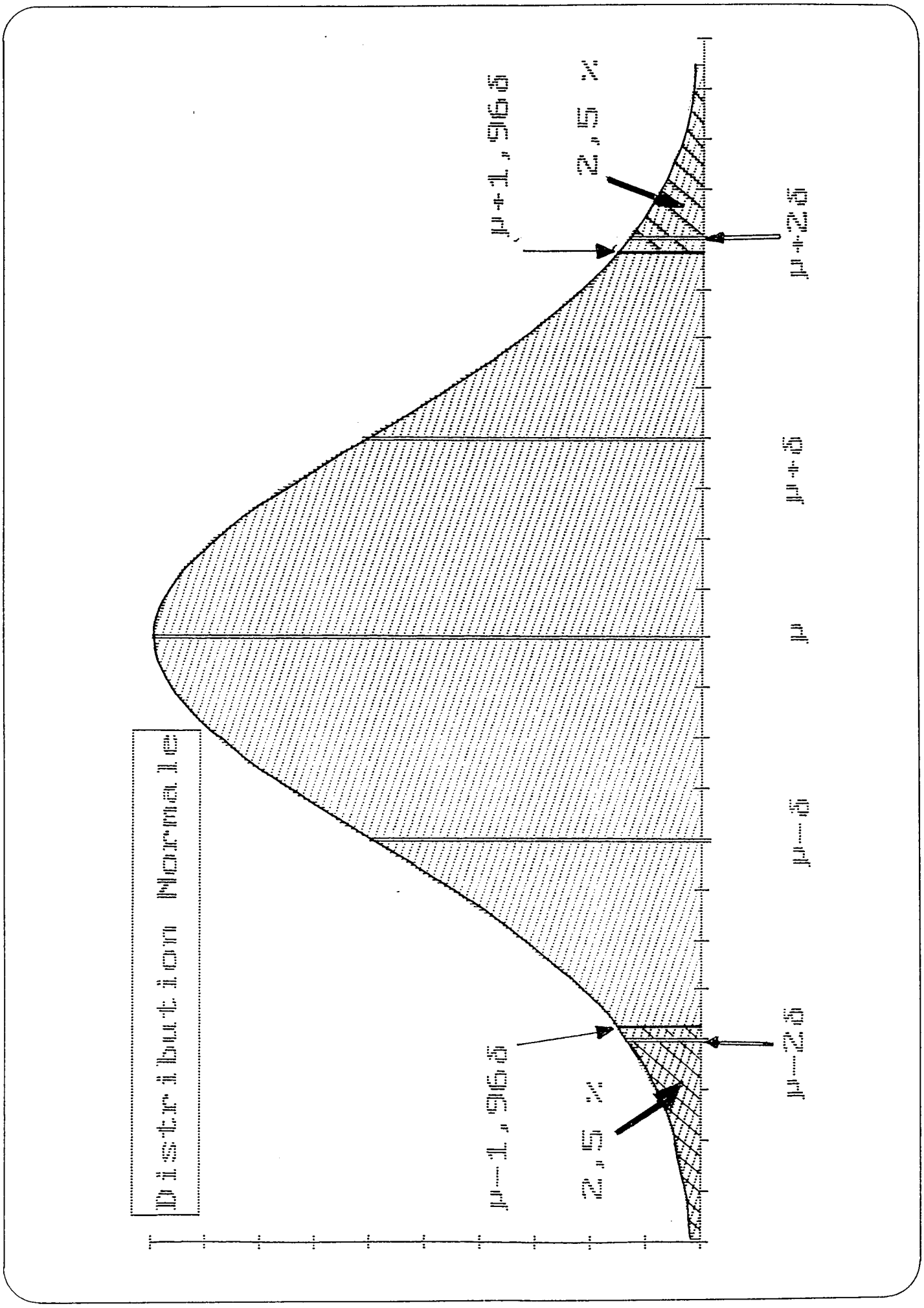
ou :

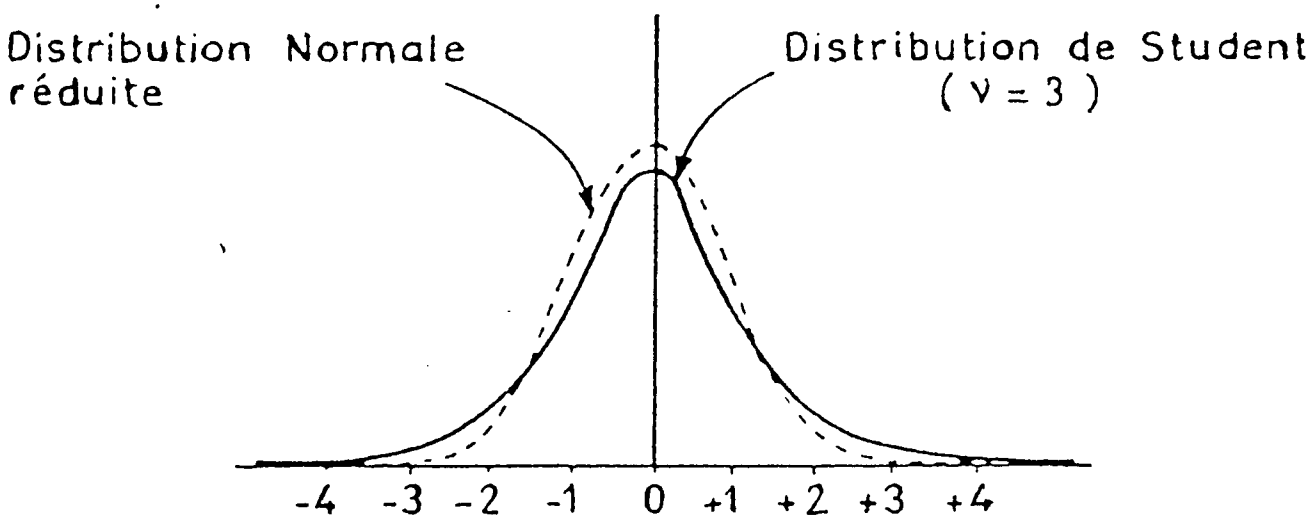
t	statistique de Student
m_1 et m_2	moyenne par lot
Sd	écart-type de la différence
S	écart-type de la population supposée unique (H_o).
n_1 et n_2	effectifs par lot.
S_1 et S_2	écart-type par lot.

Si $n_1 = n_2$

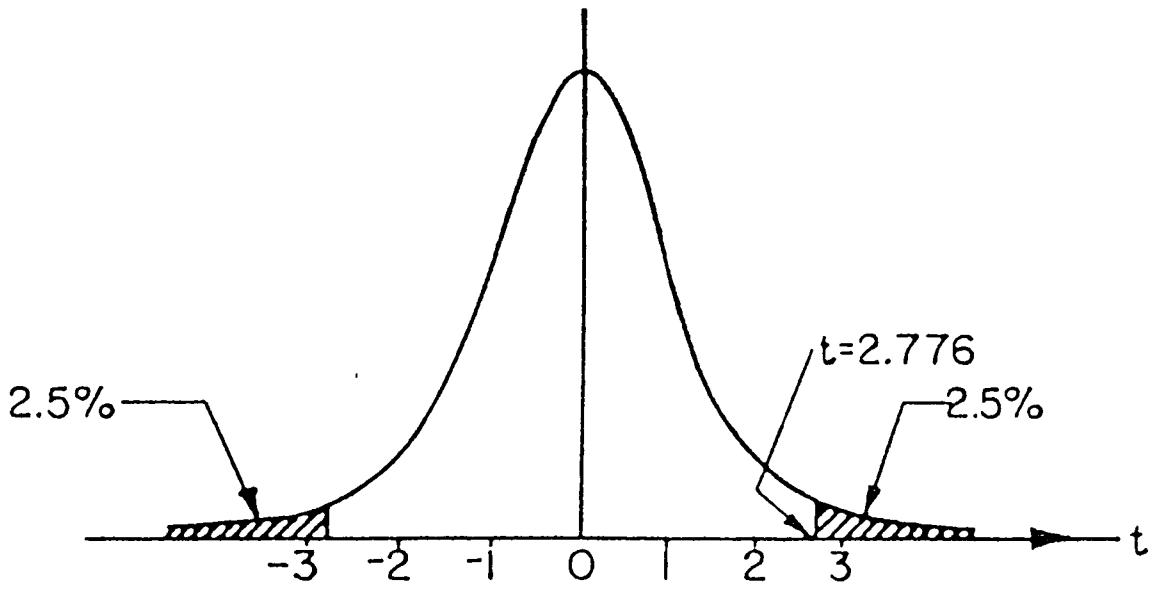
$$Sd = \sqrt{\frac{S_1^2 + S_2^2}{n}}$$

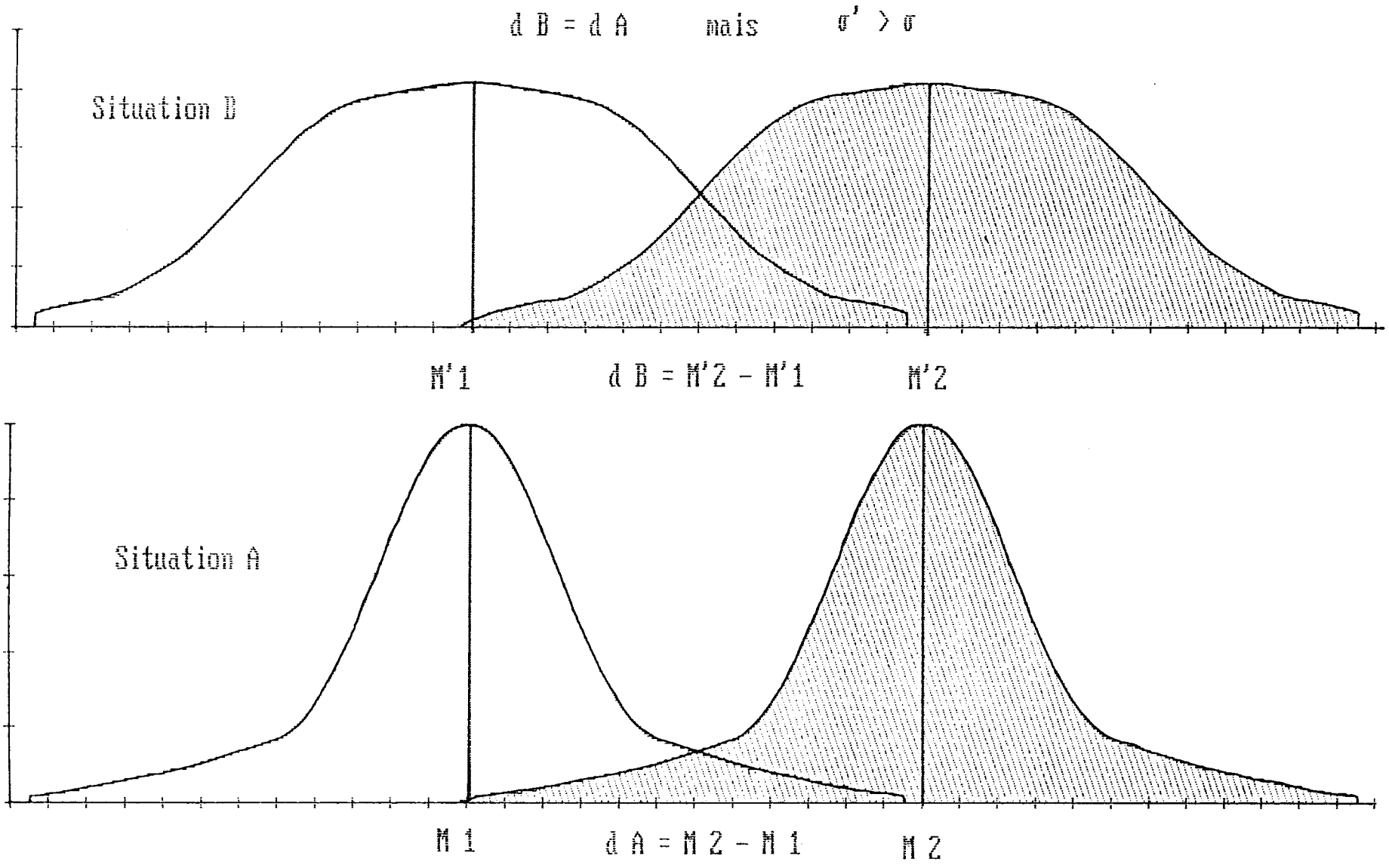
$$t = (m_1 - m_2)\,\frac{\sqrt{n}}{\sqrt{S_1^2 + S_2^2}}$$


Distribution Normale
μ+1,96δ
2,5 %
μ+2δ
μ+δ
μ
μ−δ
μ−2δ
μ−1,96δ
2,5 %



Distribution du *t* de Student et distribution normale réduite (d'après CRAMER, *Mathematical Methods of Statistics*).




d B = d A mais σ' > σ
Situation B
M'1
d B = M'2 - M'1
M'2
Situation A
M 1
d A = M 2 - M 1
M 2

Homogénéité de la Variance
μ
μ+α

LES CONDITIONS D'UTILISATION DU TEST T DE STUDENT

1. Normalité de la distribution de l'erreur expérimentale :

   En d'autres termes les résultats que l'on obtiendrait, en l'absence de tout traitement, doivent être dispersés suivant la loi normale.

2. Homogénéïté de la variance ou invariance de la variance ou homoscédasticité)

   Il faut que tous les groupes expérimentaux aient la même variance

3. Indépendance des erreurs

   Il faut que les variations non contrôlées n'aient pas de liaison entre elles, donc qu'elles se produisent strictement au hasard.

4. Additivité des effets des facteurs

   $X = u + t + b + e$ u : moyenne générale
   t : effet dû au traitement
   b : effet dû au bloc
   e : effets non contrôlés ou erreur expérimentale