

INTRODUCTION

En paléontologie, les exigences actuelles de la détermination spécifique sont celles d'une discrimination toujours plus précise et plus fine, exercée, non plus sur des individus isolés, mais sur des collections abondantes. Or, à la base de ce travail les éléments originaux des genres ou espèces choisis par de nombreux chercheurs sont souvent d'une incroyable déficience. Ces types ou holotypes ont, pour la plupart, été sélectionnés très anciennement, souvent d'après le plus bel échantillon et dans l'ignorance complète des formes moyennes correspondant à chaque phase de la croissance, ainsi que des formes déviées par l'action du biotope. Aussi de nombreux spécimens types sont-ils devenus insuffisants pour les études comparatives et descriptives. De plus beaucoup d'entre eux, éminemment précieux pourtant, ont été dispersés, soit au cours de leur pérégrination dans les institutions, soit par la négligence des collectionneurs.

A côté de ces indigences, dues au facteur humain, il en est d'autres, inhérentes à l'objet lui-même et dues à son polymorphisme et à sa fossilisation.

Les Pélécy-podes des faunes limniques ou dulcicoles des terrains houillers offrent un exemple magistral de la faible valeur de leurs types utilisés jusqu'à présent. Ils ont été prématurément choisis au sein des collections numériquement très insuffisantes et non véritablement représentatives des espèces. Cette situation a créé de multiples confusions dans les travaux paléontologiques par l'établissement de diagnoses hâtives et de déterminations approchées, qui allongent et remanient sans cesse la liste des synonymies. Il faut ajouter qu'une iconographie défectueuse complique encore la tâche du systématicien.

Ces coquilles de Pélécy-podes sont, de près, toutes semblables et, de très près, toutes différentes par la variabilité de leurs formes résultant des actions du milieu. Leur étude comparative par le simple examen visuel, ayant pour but leur classement en groupes, soit leur description morphologique, présente des insuffisances et est, au surplus, fallacieuse. Aussi les nécessités de se faire bien comprendre ont-elles amené déjà quelques chercheurs à traduire les résultats de leurs observations par des mesures.

Allant plus loin, on peut penser que l'abstraction mathématique pourrait fournir aux formes envisagées un cadre, des limites — puisqu'il n'apparaît pas de limites biologiques reconnaissables — qui soient à la fois larges et précises et suffisamment claires pour contenir l'espèce et ses variétés. Dans l'utilisation des nombres à la définition de l'espèce, rares seraient sans doute les cas où l'on pourrait envisager la « géométrisation » de sa forme et il est certain que celle-ci ne devrait être appliquée que là où les bases morphologiques de la détermination conchyliologique sont peu nombreuses, simples et mesurables avec précision.

Cette voie apparaît, dans l'avenir, riche de possibilités pour le classement systématique de ces fossiles, quand on envisage les travaux, récents ou en cours, sur la définition des coquilles par l'établissement de leur équation surfaciale basée sur la courbure du bord libre [L. LISON ⁽¹⁾].

Au préalable et présentement il est nécessaire d'accroître d'abord considérablement les collections de fossiles des espèces limniques ou dulcicoles de l'Anthracolithique; ensuite, d'opérer un classement provisoire et très prudent de ces formes sur les bases des caractères morphologiques simples et des caractères régionaux et stratigraphiques. C'est dans ce sens qu'a été conçue cette « Étude biométrique des *Anthraconauta* du houiller de la Belgique », dont la première partie est consacrée à l'examen comparatif d'*Anthraconauta belgica* HIND et d'*Anthraconauta minima* (HIND, non LUDWIG), sur les bases morphologiques usuelles, mais à l'aide de méthodes mathématiques et statistiques bien définies.

⁽¹⁾ LISON, L., 1942.

Étude biométrique des *Anthraconauta* du Houiller de la Belgique

(PREMIÈRE PARTIE)

CHAPITRE I.

§ 1. Place des mollusques limniques ou dulcicoles du Terrain houiller dans la classification zoologique systématique.

Il n'est pas sans intérêt, dans une étude sur les *Anthraconauta*, de passer en revue l'opinion des malacologistes et des paléontologistes sur la place occupée par les mollusques limniques ou dulcicoles de l'Anthracolithique dans l'une ou l'autre classification. Très peu nombreux sont ceux qui, depuis W. HIND, ont fait œuvre de systématicien en étudiant de près ce problème, à propos duquel on trouve plus de redites que de recherches originales.

Trois courants d'opinion ont cours. Le premier classe les genres *Anthracomya*, *Carbonicola* et *Naiadites* dans la famille des Unionidae. Le second, qui a rallié de nombreux avis, rapporte ces genres, ou les deux premiers de ceux-ci, aux Cardiniidae. Le troisième range ces genres dans une famille propre, tout en reconnaissant la diversité des affinités qui y règne.

M. NEUMAYR⁽²⁾ ⁽³⁾, dont la classification est basée sur les caractères de la coquille, principalement ceux de la charnière, a rangé *Anthracomya* SALTER dans les Unionidae et rapproche ces dernières des Trigonidae, alors que H. DOUVILLE⁽⁴⁾, venant après lui, faisait des *Anthracosia* des Unionidae vraies, issues de Taxodontes. Pour W. HIND⁽⁵⁾, *Anthracomya* SALTER et *Carbonicola* M'COY sont à

⁽²⁾ ⁽³⁾ NEUMAYR, M., 1893 et 1891, pp. 412 et 719.

⁽⁴⁾ DOUVILLE, H., 1912, p. 441.

⁽⁵⁾ HIND, W., 1894, pp. 12-38 et suiv.

placer dans la famille Unionidae, tout en notant que les *Anthracomya* possèdent un ensemble de caractères qui les rapprochent plus des Mytilidae que les *Carbonicola*, lesquelles seraient intermédiaires entre les premières et les Unios modernes. Quant aux *Naiadites*, rapportées tour à tour par les auteurs aux *Modiola*, aux *Mytilus*, aux *Dreissensia* (par LUDWIG) et aux *Myalina* (par BARROIS), HIND les rapporte aux Mytilidae.

DE KONING et DE RYCKHOLT ont considéré les « Unios du Houiller » comme affines du genre *Cardinia* d'AGASSIZ. K. ZITTEL, G. STEINMAN et L. DODERLEIN ⁽⁶⁾ ont rangé *Anthracosia* KING dans la famille des Cardiniidae, ordre des Hétérodontes. Dans la classification de W. H. DALL ⁽⁷⁾, on trouve encore *Carbonicola* M'COY, *Anthracosia* KING et *Naiadites* DAWSON rassemblés dans la famille des Cardiniidae de ZITTEL, superfamille des Naiadaceae, ordre des Schizodontes. C'est à cette classification que se sont ralliés la plupart des paléontologues anglo-saxons, entre autres J. HENDERSON ⁽⁸⁾, et, en Belgique, M. F. DEMANET ⁽⁹⁾. P. PELSENEER ⁽¹⁰⁾, dans sa classification zoologique basée sur les branchies, a maintenu le genre *Anthracosia* KING dans la famille des Cardiniidae de ZITTEL, mais a rangé celle-ci dans le sous-ordre des Submytilacea, ordre des Eulamellibranches. Il a classé le genre *Anthracopectera* SALTER dans la famille des Myalinidae, sous-ordre des Pectinaceae, qu'il distinguait des Mytilaceae, ces deux derniers sous-ordres faisant partie des Filibranchiés. F. BERNARD ⁽¹¹⁾ a constaté, sans prendre position, que les *Anthracosia* KING et les *Carbonicola* M'COY sont rapportées tantôt aux Cardiniidae, tantôt aux Unionidae, mais que, dans l'un ou l'autre cas, elles seraient des formes ancestrales communes aux deux groupes, comme le signalaient POHLIG et ZITTEL.

Parmi les chercheurs partisans de réunir les Lamellibranches limniques ou dulcicoles du Houiller en une famille propre, il faut citer W. AMALIZKY ⁽¹²⁾, qui groupa dans la famille des Anthracosiidae (ou Anthracosinae) les genres *Carbonicola* M'COY, *Anthracosia* KING, *Palaeomutela* AMALIZKY, *Oligodon* AMALIZKY et *Naiadites* DAWSON. Pourtant, il reconnaissait que les deux premiers genres sont des formes hétérodontes et que le genre *Naiadites* est une forme dépourvue de charnière. Pour M. P. PRUVOST ⁽¹³⁾, *Carbonicola* M'COY et *Anthracomya* SALTER doivent être rangés dans la famille des Anthracosiidae d'AMALIZKY, et *Naiadites*

⁽⁶⁾ STEINMAN, G. et DODERLEIN, L., 1890, p. 265.

⁽⁷⁾ DALL, W. H. in ZITTEL, K. and EASTMAN, R., 1927, pp. 451-452.

⁽⁸⁾ HENDERSON, J., 1935, p. 291.

⁽⁹⁾ DEMANET, F., 1938, pp. 139-145.

⁽¹⁰⁾ PELSENEER, P., 1906, pp. 262-267.

⁽¹¹⁾ BERNARD, F., 1895, p. 581.

⁽¹²⁾ AMALIZKY, W., 1892, pp. 136-142.

⁽¹³⁾ PRUVOST, P., 1930, pp. 227-247.

DAWSON et *Anthraconauta* PRUVOST, dans la famille des Mytilidae de LAMARCK. L'homogénéité de la famille des Anthracosiidae a été reconnue encore par H. PAUL ⁽¹⁴⁾. Une opinion récente de M. A. E. TRUEMAN place les *Carbonicola* du groupe de *C. aquilina* SOWERBY dans la famille des Anthracosiidae ⁽¹⁵⁾.

§ 2. Phyllogénèse des *Anthraconauta*.

Les opinions émises ces dernières années sur l'origine phyllétique des *Anthraconauta* traduisent une grande diversité de conception. M. D. FEDOTOV ⁽¹⁶⁾ rattache les grandes *Anthraconauta* du bassin de Kusnetzki aux *Anthracomya* à faible obliquité. De son côté, M. H. WEHRLI ⁽¹⁷⁾ a exprimé l'idée d'une descendance phyllétique issue des *Naiadites* ⁽¹⁸⁾, opinion qui a subi la critique de MM. E. DIX et A. E. TRUEMAN ⁽¹⁹⁾. Ces deux thèses ont été discutées récemment par M. J. WEIR ⁽²⁰⁾. Ce paléontologue a signalé la possibilité de trouver une origine phyllétique des *Anthraconauta* dans le groupe des *Posidonia* de la famille des Pteriidae (Aviculidae). Il estime cependant possible, à l'origine, une séparation diphyllétique des *Anthraconauta* et des *Naiadites* (provenant les unes et les autres d'un stock commun au stade « *Modiola* ») rapidement différenciées et stabilisées en une morphologie distincte ⁽²¹⁾.

Cette suggestion avait été faite auparavant par M. N. D. NEWELL ⁽²²⁾, qui a étudié d'une façon approfondie le problème des relations entre Myalinidae et Mytilidae du Paléozoïque de l'Amérique du Nord. Après avoir constaté que la tendance à l'obliquité des valves et le déplacement antérieur de l'umbo n'est pas une tendance particulière aux Myalinidae, mais se retrouve également chez les Mytilidae, les Ambonichiidae et les Pteriidae ⁽²³⁾, il conclut à un cas d'homéomorphisme, non nécessairement phylogénétique, entre les formes de ces groupes. De plus, il reconnaît que les Mytilidae primitives du Paléozoïque présentent une expression « modiolloïde » très nette, ce qui l'incite à classer la majorité de leurs formes ayant un aspect « mytiloïde » ou « modiloïde » parmi les Mytilacea, sous quelques réserves cependant.

Il lui apparaît que, par la forme et la structure de la charnière, le genre

⁽¹⁴⁾ PAUL, H., 1941, pp. 276-277.

⁽¹⁵⁾ TRUEMAN, A. E., 1941, communication écrite; voir aussi : TRUEMAN, A. E., 1946.

⁽¹⁶⁾ FEDOTOV, D., 1937.

⁽¹⁷⁾ WEHRLI, H., 1933, pp. 133-134.

⁽¹⁸⁾ WEIR, J., 1944, pp. 316 et suiv.

⁽¹⁹⁾ DIX, E. et TRUEMAN, A. E., 1931, pp. 192-193.

⁽²⁰⁾ WEIR, J., 1944, p. 36.

⁽²¹⁾ WEIR, J., 1944, p. 321.

⁽²²⁾ NEWELL, N. D., 1940.

⁽²³⁾ NEWELL, N. D., 1940, p. 8.

Modiolopsis HALL, déjà présent à l'Ordovicien, pourrait être vraisemblablement l'ancêtre de tous les Mytilidae. Les espèces de ce genre sont proches des *Volsella* (*Modiolus*) actuelles ⁽²⁴⁾.

Cette opinion, à laquelle nous nous rallions, reprend des idées émises anciennement par F. BERNARD ⁽²⁵⁾, qui voyait dans les *Modiola* et les *Modiolopsis* d'ULRICH des époques précambriennes, les stades divers du développement des Mytilidae.

§ 3. Caractères généraux des *Anthraconauta*.

C'est à M. P. PRUVOST que revient la création du genre *Anthraconauta*, qu'il rattache à la famille des Mytilidae de LAMARCK en séparant les espèces de ce genre du groupe générique désigné sous le nom d'*Anthracomya*, famille des Anthracosiidae d'AMALIZKY. Cet auteur a exposé longuement les raisons qui l'ont amené à établir cette coupure, dans un travail sur la faune continentale du Westphalien de la Belgique ⁽²⁶⁾, dont voici extraite la diagnose du genre : « Valves de contour subovale, allongées obliquement (angle d'environ 30°) par rapport au bord cardinal, toujours plus court que la plus grande longueur. Large échancrure byssale peu profonde sur le bord antéro-ventral, séparant la valve en une partie antérieure arrondie en oreille et une postérieure de beaucoup plus importante. Une carène mousse rectiligne, partant du crochet, forme la ligne de plus grande longueur. Caractères internes inconnus. Test de structure prismatique, recouvert d'un périostracum épais, souvent ridé, portant des stries d'accroissement concentriques irrégulièrement espacées. Formes vivant associées, en colonies, flottant fixées par leur byssus à un support » ⁽²⁷⁾.

A l'époque de la publication de l'ouvrage, l'auteur avait reconnu l'existence des cinq espèces suivantes : *A. Phillipsi* (WILLIAMSON), génotype, *A. minima* (LUDWIG), *A. laevis* (DAWSON), *A. belgica* (HIND) et *A. stephaniensis* PRUVOST. Depuis, la liste s'est augmentée de : *A. tenuis* DAVIES et TRUEMAN ⁽²⁸⁾, *A. Wrighty* DIX et TRUEMAN ⁽²⁹⁾, *A. anthracomyoides* FEDOTOV ⁽³⁰⁾, *A. krochalovskiensis* FEDOTOV ⁽³¹⁾, *A. tenuistriata* FEDOTOV ⁽³²⁾, *A. iljinskiensis* FEDOTOV ⁽³³⁾, *A. pseudophilipsi* FEDOTOV ⁽³⁴⁾, *A. kemeroviensis* FEDOTOV ⁽³⁵⁾, *A. gigantea* RAGOZIN ⁽³⁶⁾,

⁽²⁴⁾ NEWELL, N. D., 1940, pp. 35-36.

⁽²⁵⁾ BERNARD, F., 1898, pp. 196-197.

⁽²⁶⁾ PRUVOST, P., 1930, pp. 223-225.

⁽²⁷⁾ PRUVOST, P., 1930, p. 247.

⁽²⁸⁾ DAVIES, J. H. and TRUEMAN, A. E., 1927, p. 235.

⁽²⁹⁾ DIX, E. and TRUEMAN, A. E., 1931, p. 184.

⁽³⁰⁾ ⁽³¹⁾ ⁽³²⁾ FEDOTOV, D. M., 1938, pp. 249-250.

⁽³³⁾ ⁽³⁴⁾ ⁽³⁵⁾ FEDOTOV, D. M., 1937, pp. 103-104.

⁽³⁶⁾ RAGOZIN, U., 1933, pp. 56-63.

A. sibirica RAGOZIN ⁽³⁷⁾, *A. minuta* RAGOZIN ⁽³⁸⁾, *A. longa* RAGOZIN ⁽³⁹⁾, *A. flagitiosa* RAGOZIN ⁽⁴⁰⁾, *A. pungabica* REED ⁽⁴¹⁾, toutes espèces peu connues encore. Récemment, le terrain houiller du Lancashire a permis à M. DEWAR de distinguer les espèces et variétés suivantes : *A. minima* (HIND, non LUDWIG), var. *obliqua*, *A. trapeziforma* DEWAR, *A. trapeziforma* DEWAR var. *tenuides* DEWAR, *A. subovata* DEWAR, *A. subovata* var. *candela* DEWAR, et de redécrire *A. minima* (HIND, non LUDWIG) et *A. scotica* ETHERIDGE ⁽⁴²⁾.

Il faut noter que quelques auteurs, H. WEHRLI et P. KUKUK, continuent de ranger les *Anthraconauta* distinguées par M. P. PRUVOST parmi les *Anthracomya*.

§ 4. Spécificité d'*Anthraconauta belgica* HIND.

Parmi les espèces distinguées par M. P. PRUVOST, l'*Anthraconauta belgica* HIND est une des moins connues, parce que de rencontre très rare. Tous les gisements qui en ont livré des exemplaires sont situés dans le bassin houiller de Liège, en Belgique. L'examen des quelques échantillons que possèdent les collections de nos institutions a éveillé des doutes sérieux sur la spécificité de l'espèce, ce qui nous a conduits à entreprendre une revision critique du sujet.

Voici la diagnose originelle établie par W. HIND : « Coquille petite, ovalemment subcirculaire; l'extrémité antérieure est courte, profonde, avec un bord régulièrement courbé. Le bord inférieur est presque droit et descend rapidement. L'angle postérieur, inférieur, est très dilaté et brusquement arrondi. Le bord postérieur, court, presque droit, est tronqué obliquement à partir du haut vers le bas. La charnière est droite, courte et postérieurement élevée. L'angle postérieur, inférieur, est très dilaté et brusquement arrondi. Le bord postérieur, court, presque droit, est tronqué obliquement à partir du haut vers le bas. La charnière est droite, courte et postérieurement élevée. L'angle postérieur est bien marqué et obtus. Les crochets, petits et vaguement marqués (inconspicuous), sont placés dans le tiers antérieur de la valve. La coquille est très dilatée et aplatie postérieurement avec une dépression marquée le long de la pente dorsale. La valve est ornée de stries d'accroissement concentriques, fines et serrées. Périostracum ridé » ⁽⁴³⁾. Il ressort de cette description que les caractères principaux de l'espèce seraient : la forme orbiculaire plus arrondie que celle des autres *Anthraconauta*, la région postérieure moins développée et le bord dorsal moins oblique. D'après les observations de M. P. PRUVOST, « la position subcentrale de

⁽³⁷⁾ RAGOZIN, U., 1933, pp. 56-63.

⁽³⁸⁾ ⁽³⁹⁾ ⁽⁴⁰⁾ RAGOZIN, U., 1935, pp. 57-74.

⁽⁴¹⁾ REED, F. R. G., 1939.

⁽⁴²⁾ DEWAR, W., 1939, pp. 54-61.

⁽⁴³⁾ HIND, W., 1911, p. 17.

son crochet et par suite le développement de l'aile antérieure en font un type distinct des autres »⁽⁴⁴⁾. Ces caractères spécifiques ont également été reconnus par M. F. DEMANET⁽⁴⁵⁾ ⁽⁴⁶⁾.

§ 5. Nécessité d'une étude biométrique.

Un examen attentif et soigné des spécimens types décrits et figurés par W. HIND⁽⁴⁷⁾ et reproduits ailleurs par M. F. DEMANET⁽⁴⁸⁾ ⁽⁴⁹⁾ révèle ce qui suit :

Le spécimen des figures 1 et 1a, reproduites par W. HIND, est le moulage d'une empreinte incomplète dans la région de l'angle postéro-inférieur, où il manque 1 à 2 mm. pour présenter un contour valvaire normal, parallèle aux stries concentriques d'accroissement. De plus, dans la région cardinale, le test a été plié et retroussé dans la plaquette schisteuse, ce qui a eu pour conséquence une déformation de la ligne de la charnière ne permettant pas une bonne appréciation de la position de l'umbo. La valve est fortement aplatie, fait qui contribue à lui donner un aspect plus arrondi.

Le spécimen des figures 2 et 2a, du même auteur, a subi aussi un aplatissement déformateur du contour valvaire. Au demeurant, le meilleur de ceux reproduits par ces chercheurs est l'exemplaire de la figure 3. Il se présente, en relief, non déformé, avec un contour valvaire net et bien délimité sur les deux tiers de son pourtour; seule la région du crochet est imprécise. C'est sans conteste une empreinte d'*Anthraconauta*.

D'autres spécimens, catalogués sous le nom d'*A. belgica*, les uns récoltés par M. F. DEMANET aux gisements d'Angleur et d'Andenne; les autres, des diverses provenances énumérées au § 6, ont été examinés. Une étude soignée de tous ces exemplaires nous a incités à mettre en doute la spécificité de *A. belgica*, parce que le caractère distinctif principal, l'orbicularité des valves, apparaît le plus souvent n'être qu'accidentel, comme un effet de leur écrasement. Les exemplaires rangés sous ce nom ne nous ont pas paru, à l'examen visuel, différents de ceux définis comme *A. minima*. Une étude biométrique, entreprise sur un matériel beaucoup plus abondant que celui récolté jusqu'alors, s'imposait pour confirmer ou infirmer cette observation.

⁽⁴⁴⁾ PRUVOST, P., 1930, p. 247.

⁽⁴⁵⁾ DEMANET, F., 1938, p. 144.

⁽⁴⁶⁾ DEMANET, F., 1941, p. 261.

⁽⁴⁷⁾ HIND, W., 1911, pl. I, fig. 1-3.

⁽⁴⁸⁾ DEMANET, F., 1938, pl. 119, fig. 11 à 13.

⁽⁴⁹⁾ DEMANET, F., 1941, pl. 15, fig. 38. 39.

§ 6. Position stratigraphique des gisements et provenances
du matériel étudié.

Les gisements connus appartiennent tous aux terrains des âges namuriens (*Nm*) et westphalien inférieur (*Wn1a*). Voici les localités d'origine des exemplaires examinés et qui représentent pratiquement l'ensemble des spécimens existant dans les collections de Belgique :

ÉTAGE NAMURIEN. — ASSISE D'ANDENNE (*Nm2*).

Pl. Chenée :

1. Angleur, en face de la fabrique de produits chimiques. [Coll. Mus. roy. Hist. nat. Belg., Bruxelles, coll. X. STAINIER, 1891-1901, I.G. 11.312, dét. X. STAINIER : *Modiola carlottae* ROEMER; rév. F. DEMANET, 1939 : *A. belgica* (HIND).]

2. Kinkempois (Angleur), gisement non précisé. (Coll. Université de Liège, coll. géologique, n° 92, dét. X... : *Mytilus carlottae* ROEMER.)

3. Angleur, gisement non précisé. [Coll. Mus. roy. Hist. nat. Belg., Bruxelles, anc. coll. de KONINCK et RUTOT. I.G. 8573. Coll. des spécimens figurés, dét. W. HIND, 1902 : *Anthracomya belgica* sp. nov.; rév. P. PRUVOST, 1921 : *Modiola belgica* HIND; rév. F. DEMANET, 1930 : *A. belgica* (HIND). Plusieurs spécimens types, reproduits in : HIND, W., 1911, t. VI, fig. 1-3, et in : DEMANET, F., etc., 1938, pl. 119, fig. 11-13.]

4. Angleur, à flanc de colline, à 500 m. au S.-O. de la Station. Points n°s 21.506 et 21.508. (Coll. Mus. roy. Hist. nat. Belg., Bruxelles, I.G. 12.527, coll. des spécimens figurés, dét. F. DEMANET : *A. belgica* HIND, spécimen type reproduit in : DEMANET, F., 1941, pl. XV, fig. 38 et 39.)

5. Angleur, à flanc de colline, à 180 m. au S.-E. du tunnel-abri de la rue Jean Volders. (Coll. Ass. Étude Pal. Strat. Houil., I.G. 39, 68.)

Pl. Andenne :

Andenne, bois de Stud, carrière. Point n° 19. (Coll. Mus. roy. Hist. nat. Belg., I.G. 11.516, dét. F. DEMANET, 1939 : *A. belgica* HIND. Coll. Ass. Étude Pal. Strat. Houil., I.G. n°s 37, 67.)

Pl. Herve :

Val-Dieu, ferme de la Mouderie. Gisements n°s VIII et XVIII. [Coll. Ass. Étude Pal. Strat. Houil., I.G. n° 70, dét. CH. ANCIEN et W. VAN LECKWIJCK : *A. belgica* (HIND).]

Charbonnage de Forte-Taille. Sondage n° 26 (4) à Jamioulx. A 244, 246, 247 et 248,50 m. de profondeur, niveaux n°s 11 et 12 de la coupe. [Coll. Mus. roy. Hist. nat. Belg., Bruxelles, coll. X. STAINIER, I.G. n° 11.312, dét. X... : *Posidoniella* sp., rév. F. DEMANET, 1939 : *A. belgica* (HIND).]

ÉTAGE WESTPHALIEN. — ASSISE DE CHÂTELET. (*Wn1a, Wn1b*).

Charbonnage de Courcelles-Nord, puits n° 8, sondage intérieur, de 361 à 363 m. de profondeur, niveau n° 13, n° 157 de la coupe. [Coll. Mus. roy. Hist. nat. Belg., I.G. 14.632, dét. P. PRUVOST, 1930 : *A. belgica* (HIND).]

Charbonnage d'Amercœur, siège : Naye-à-Bois, étage de 790 m., à 96,80 m. au Nord du puits, à 50 m. sur veine Sainte-Barbe de Floriffoux. [Coll. Mus. roy. Hist. nat. Belg., I.G. 14.632, dét. P. PRUVOST, 1930 : *A. belgica* (HIND).]

CHAPITRE II.

INTRODUCTION.

Les grandes lignes d'une étude biométrique de populations d'individus de provenances différentes et soupçonnés appartenir à une même espèce sont les suivantes :

A. — Choix des caractères morphologiques à mesurer. Grandeur de la précision des mesures.

B. — Choix des méthodes de la statistique mathématique à appliquer à l'étude biométrique.

C. — Mensuration des individus de chaque population.

Pour chaque population :

D. — Étude de la corrélation des caractères.

E. — Étude des indices caractéristiques (rapports entre les mesures des caractères des individus) et de leur distribution de fréquences.

F. — Conclusions biologiques relatives à la « variabilité individuelle » et aux « variations spécifiques ».

G. — Comparaison des indices types des différentes populations pour la recherche d'une conspécificité éventuelle de leurs éléments.

Dans le présent chapitre, nous exposerons les trois premiers points du schéma précité en développant particulièrement l'exposé de la méthode employée. Il ne convient pas ici de s'étendre sur toutes les méthodes mathématiques qui sont susceptibles d'être utilisées en biométrie. Signalons toutefois la grande indigence dont font preuve nombre de biologistes qui n'utilisent pas pour leurs recherches toutes les ressources de l'analyse mathématique et de la statistique, ou qui mettent en pratique, sans discernement, des recettes toutes

ÉTAGE WESTPHALIEN. — ASSISE DE CHÂTELET. (*Wn1a, Wn1b*).

Charbonnage de Courcelles-Nord, puits n° 8, sondage intérieur, de 361 à 363 m. de profondeur, niveau n° 13, n° 157 de la coupe. [Coll. Mus. roy. Hist. nat. Belg., I.G. 14.632, dét. P. PRUVOST, 1930 : *A. belgica* (HIND).]

Charbonnage d'Amercœur, siège : Naye-à-Bois, étage de 790 m., à 96,80 m. au Nord du puits, à 50 m. sur veine Sainte-Barbe de Floriffoux. [Coll. Mus. roy. Hist. nat. Belg., I.G. 14.632, dét. P. PRUVOST, 1930 : *A. belgica* (HIND).]

CHAPITRE II.

INTRODUCTION.

Les grandes lignes d'une étude biométrique de populations d'individus de provenances différentes et soupçonnés appartenir à une même espèce sont les suivantes :

A. — Choix des caractères morphologiques à mesurer. Grandeur de la précision des mesures.

B. — Choix des méthodes de la statistique mathématique à appliquer à l'étude biométrique.

C. — Mensuration des individus de chaque population.

Pour chaque population :

D. — Étude de la corrélation des caractères.

E. — Étude des indices caractéristiques (rapports entre les mesures des caractères des individus) et de leur distribution de fréquences.

F. — Conclusions biologiques relatives à la « variabilité individuelle » et aux « variations spécifiques ».

G. — Comparaison des indices types des différentes populations pour la recherche d'une conspécificité éventuelle de leurs éléments.

Dans le présent chapitre, nous exposerons les trois premiers points du schéma précité en développant particulièrement l'exposé de la méthode employée. Il ne convient pas ici de s'étendre sur toutes les méthodes mathématiques qui sont susceptibles d'être utilisées en biométrie. Signalons toutefois la grande indigence dont font preuve nombre de biologistes qui n'utilisent pas pour leurs recherches toutes les ressources de l'analyse mathématique et de la statistique, ou qui mettent en pratique, sans discernement, des recettes toutes

faites, telle la célèbre loi normale de GAUSS qui n'est au demeurant qu'un cas particulier de la distribution des fréquences. Aussi développerons-nous les notions mathématiques utilisées dans cette étude pour qu'aucun doute ne subsiste ni sur l'opportunité d'adopter certaines méthodes plutôt que des formules toutes faites, ni sur la nécessité de pousser dans toute sa rigueur et dans le détail l'analyse statistique et mathématique des séries.

A. — CHOIX DES CARACTÈRES MORPHOLOGIQUES A MESURER.

GRANDEUR DE LA PRECISION DES MESURES.

CARACTÈRE GÉNÉRAL DES MENSURATIONS EFFECTUÉES. — Les problèmes de caractérisation et d'ajustement de séries statistiques que s'efforce de résoudre la statistique mathématique, et que nous utiliserons, portent sur des variables dénommées : variables statistiques. Ces variables ne sont pas autre chose que l'expression quantitative des caractères observés et mesurés. Ces caractères sont susceptibles de varier dans certaines limites de deux façons : ou bien la variable statistique ne peut prendre qu'un nombre fini de valeurs formant une suite discrète ou discontinue : x_1, x_2, \dots, x_n ; ou bien la variable est une variable continue, pouvant prendre toutes les valeurs comprises entre les limites extrêmes de la variation.

Dans beaucoup de cas un grand nombre de caractères biométriques, et notamment les tailles d'individus, les mesures de longueurs et d'angles caractéristiques des coquilles de lamellibranches, sont des variables statistiques continues.

GRANDEURS MESURÉES. — La figure 1 représente les grandeurs mesurées sur les valves du type de l'*Anthraconauta* et leur notation par MM. P. PRUVOST, J. H. DAVIES et A. E. TRUEMAN et par nous-mêmes. En conclusion de cette étude, nous montrerons que certaines de ces grandeurs ne sont point caractéristiques et sont dépourvues de signification précise.

Pour P. PRUVOST ⁽⁵⁰⁾, dans la figure 1 (1), les grandeurs :

c = longueur de la ligne cardinale,

b = longueur de la carène,

$a-b$ = la plus grande largeur.

Pour J. H. DAVIES et A. E. TRUEMAN ⁽⁵¹⁾, dans la figure 1 (2), les grandeurs :

L = longueur de la valve,

H = hauteur,

A = longueur de la région antérieure.

⁽⁵⁰⁾ PRUVOST, P., 1920, fig. 1 à 3, p. 7 et fig. 8-12, p. 30.

⁽⁵¹⁾ DAVIES, J. H. et TRUEMAN, A. E., 1927, p. 212.

Au cours de leur emploi dans une analyse biométrique, les mesures H, A et T (épaisseur) sont souvent exprimées comme pourcentage par rapport à la longueur, et la caractérisation géométrique de la coquille s'établit ainsi : — L; $\frac{H}{L}$; $\frac{T}{L}$; $\frac{A}{L}$.

Pour notre étude, nous avons représenté, à la figure 1 (3), les grandeurs ayant fait l'objet de mesures; elles ne nécessitent pas de commentaires; nous avons appelé g la somme $e+c$.

Pour M. GLIBERT ⁽⁵²⁾, les grandeurs a et b , de la figure 1 (3), sont les diamètres umbono-ventral et antéro-postérieur, et le rapport $\frac{a}{b}$ l'indice des diamètres.

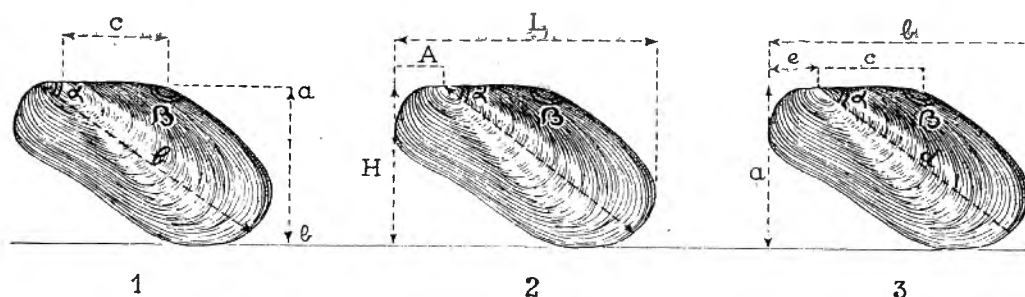


FIG. 1.

La mesure des grandeurs précitées fut effectuée, sur projection optique agrandie de l'image, au moyen d'un pied à coulisse de précision et d'un goniomètre d'application. La précision des mesures linéaires fut de $\pm 0,1$ mm. et la précision des mesures angulaires de 1 à 5°, suivant l'état du spécimen.

B. — CHOIX DES METHODES DE LA STATISTIQUE MATHÉMATIQUE A APPLIQUER A L'ÉTUDE BIOMÉTRIQUE (*).

SÉRIES STATISTIQUES A UNE VARIABLE.

§ 1. Distribution de fréquences. Représentation géométrique.

Soit x une variable statistique discontinue susceptible de prendre n valeurs distinctes x_1, x_2, \dots, x_n . Le plus souvent, en pratique, ces valeurs forment une suite en progression arithmétique, les différences entre valeurs successives étant toutes égales entre elles.

⁽⁵²⁾ GLIBERT, M., 1943, p. 5.

(*) Ouvrages généraux à consulter :

YULE, G. U. et KENDALL, M., 1945; RISSER, R. et TRAYNARD, C. E., 1933; RISSER, R., 1932; LINDER, A., 1945; VAN DEUREN, P., 1934-1935; FELDMAN, W. M., 1935; DARMOIS, G., 1941.

Un certain nombre $N > n$ d'expériences montrent que ces n valeurs sont observées respectivement $\varphi(x_1)$, $\varphi(x_2)$, ..., $\varphi(x_n)$ fois. On a évidemment

$$\varphi(x_1) + \varphi(x_2) + \dots + \varphi(x_n) = \sum_{i=1}^n \varphi(x_i) = N.$$

Les quantités $\varphi(x_i)$ et $\frac{\varphi(x_i)}{N}$ sont les fréquences absolues et les fréquences relatives ou probabilités statistiques des x_i . Le nombre N sera toujours supposé assez grand pour qu'on puisse le qualifier de « grand nombre » au sens général

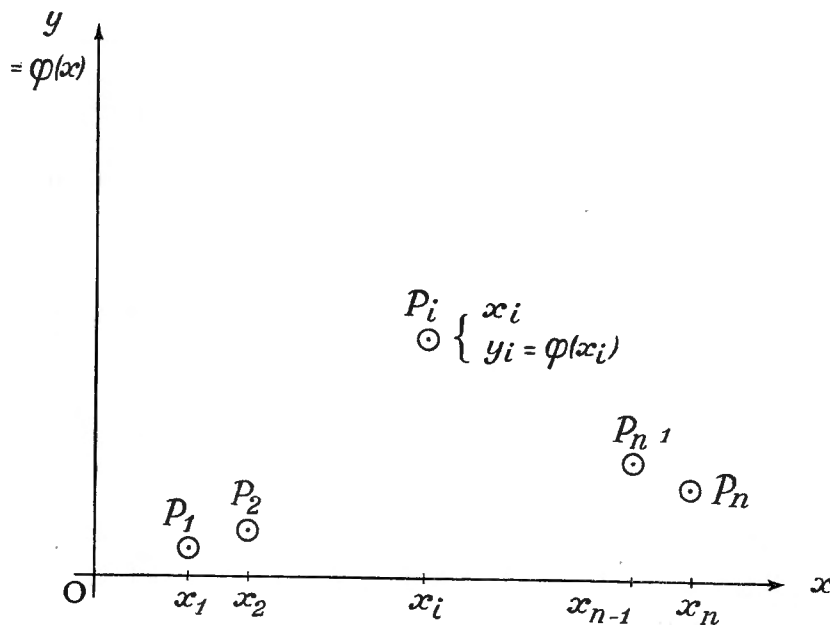


FIG. 2.

qu'on accorde à cette expression dans le vocabulaire des probabilités. Nous dirons qu'un grand nombre est un nombre tel que les conclusions de nombreuses propositions du calcul des probabilités, à caractère asymptotique, relatives à des nombres croissant au delà de toute limite, soient applicables de façon légitime, avec une erreur négligeable, mais bien connue.

On obtient ainsi une distribution des fréquences observées suivant les valeurs de la variable statistique. Cette distribution est susceptible d'être représentée géométriquement, sous forme d'un graphique à n points isolés rapportés à un système d'axes x et $y = \varphi(x)$. Ces points $P_1, P_2, \dots, P_i, P_n$ ont pour coordonnées x_i et $\varphi(x_i)$. Relier ces points par une courbe continue ou par une ligne brisée n'a pas de signification analytique, la fonction $\varphi(x)$ n'étant définie qu'en x_1, x_2, \dots, x_n (fig. 2).

Considérons à présent une variable statistique continue x , susceptible de prendre toutes les valeurs possibles entre deux limites extrêmes a et b . Soit $p(x)$

une fonction telle que $p(x_0)dx$ soit le nombre d'expériences ayant révélé une valeur de x comprise entre $x_0 - \frac{1}{2} dx$ et $x_0 + \frac{1}{2} dx$, et supposons que dx soit un infiniment petit. La fonction $p(x)$ est le taux de fréquence de la variable x .

Si N est le nombre d'expériences, on obtient, les intégrations se substituant aux sommations,

$$\int_a^b p(x) dx = N.$$

Le nombre d'expériences donnant lieu à l'observation de valeurs de x comprises entre α et β ($a \leq \alpha < \beta \leq b$) est

$$\int_a^{\beta} p(x) dx.$$

Quant à la fréquence relative ou probabilité statistique d'obtenir une valeur de x quelconque comprise entre α et β , elle s'exprime par

$$\frac{1}{N} \int_a^{\beta} p(x) dx = \frac{\int_a^{\beta} p(x) dx}{\int_a^b p(x) dx} = \int_a^{\beta} T_x dx,$$

où

$$T_x = \frac{p(x)}{\int_a^b p(x) dx}$$

est le taux de probabilité statistique.

En pratique, un relevé relatif à N expériences sur le compte d'une variable continue se fait dans l'ignorance de la fonction $p(x)$. Pour se rapprocher le mieux possible des fréquences théoriques, infiniment petites $p(x)dx$, on partagera l'intervalle (ab) en n intervalles égaux, de longueur h , centrés sur les points x_1, x_2, \dots, x_n , tels que

$$\begin{aligned} x_1 &= a + \frac{h}{2} \\ x_n &= b - \frac{h}{2} \\ x_k &= x_1 + (k-1)h. \end{aligned}$$

On notera les fréquences absolues relatives aux n intervalles ainsi définis par $\varphi(x_1)h, \varphi(x_2)h, \dots, \varphi(x_n)h$.

On définit ainsi une fonction à paliers $y = \varphi(x)$, qui, dans chaque intervalle partiel, reste constante et égale à $\varphi(x_i)$. Dans ces conditions, la fréquence $\varphi(x_i)h$ a pour mesure l'aire du rectangle ABCD (fig. 3) et la somme des aires de tous les

rectangles analogues, de même base h et de hauteurs $\varphi(x_1)$, $\varphi(x_2)$, ... $\varphi(x_n)$, est égale à N :

$$\sum_{i=1}^n \varphi(x_i) h = N.$$

La figure ainsi obtenue est appelée habituellement polygone des rectangles.

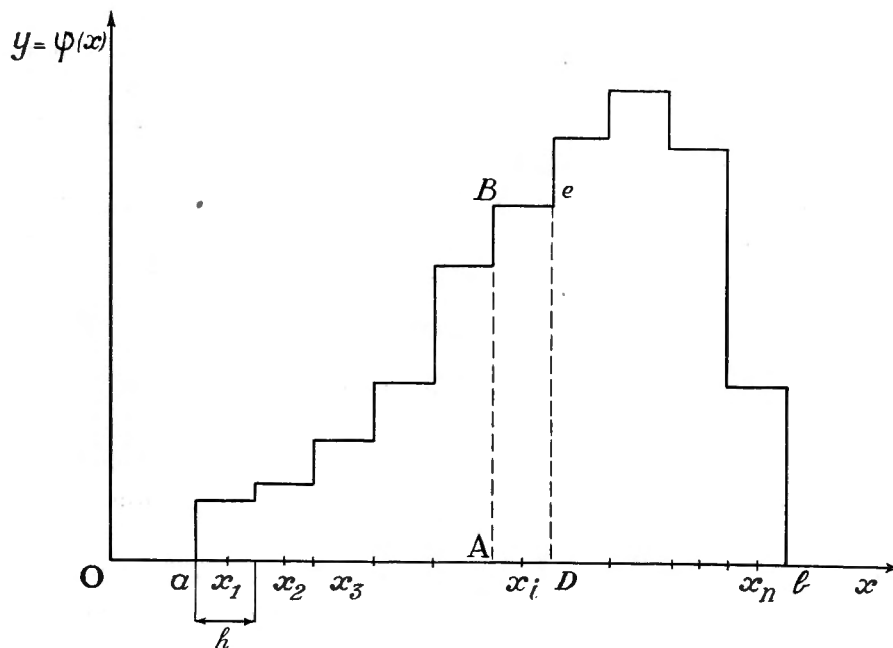


FIG. 3.

§ 2. Problème fondamental de l'ajustement des séries statistiques.

Si l'on veut tirer des conclusions, susceptibles de s'exprimer en langage mathématique, d'un relevé statistique portant sur la mesure d'un caractère jouant le rôle de la variable statistique, on ne peut se contenter d'une distribution expérimentale brute, telle que celle que nous venons de décrire. On ne peut se contenter davantage de la caractérisation d'une série statistique au moyen de quelques constantes qu'on estime significatives (moyennes, dispersion, mode, médiane, etc.). Ce sont, en effet, des résultats ou bien de l'application de formules élémentaires, ou bien de la considération d'un graphique « brut », ou bien encore d'une numérotation dans un ordre naturel de croissance ou de décroissance des valeurs expérimentales. Ces nombres sont si peu « caractéristiques » qu'il n'est guère compliqué de trouver des séries dont la représentation géométrique fait apparaître une complète dissemblance, mais qui peuvent avoir en commun certains de ces nombres calculés au moyen de formules simplistes.

Si l'on opère des mesures dans des conditions matérielles identiques sur des objets, des êtres d'un même groupe ou de même espèce, on constatera que les

lois de distribution des fréquences obtenues de plusieurs séries d'un grand nombre N d'expériences ne sont pas rigoureusement identiques. Les répartitions brutes se « ressemblent », elles sont assez voisines les unes des autres, mais on observe des écarts d'une répartition à l'autre.

Il s'ensuit, nécessairement, que les constantes statistiques relatives à chaque série de N expériences varieront de l'une à l'autre de ces séries.

Quel parti peut-on tirer, dès lors, des résultats d'une série de N expériences sur des individus pris arbitrairement parmi une collection plus nombreuse ? Quelles conclusions l'étude de cette série permet-elle d'établir, conclusions applicables à toute la collection, à toute l'espèce ? Quels sont les fondements logiques de ces conclusions et quelles en sont les limites ?

Il peut être répondu à ces questions en disant que les résultats expérimentaux bruts sont les données d'un problème dont la résolution et les conclusions sont fondées sur la théorie des probabilités. Les limites de ces conclusions dépendront directement des écarts qu'on pourra considérer avec une quasi-certitude comme ne devant jamais être dépassés.

On recherchera si la série brute étudiée ne présente pas certaines régularités qu'on pourrait traduire par une loi mathématique simple. C'est dans la recherche de cette loi mathématique que consiste le problème de l'ajustement de la série statistique à une variable. Cette loi s'exprimera au moyen d'une fonction de la variable statistique.

Deux problèmes, algébrique et analytique, consistent à déterminer une fonction $y=f(x)$ au moyen de la donnée des valeurs prises par cette fonction pour un certain nombre de valeurs de x , ou au moyen de la donnée des valeurs de l'intégrale définie de cette fonction sur un certain nombre d'intervalles contigus de la variation de x . Ainsi posés, ces problèmes sont complètement indéterminés et il convient de choisir à priori la forme des fonctions.

Or, il ne peut être question d'identifier les valeurs directement observées aux valeurs prises par la fonction inconnue, afin de la déterminer par des méthodes d'algèbre ou d'analyse mathématique. On sait, par expérience, que les valeurs observées fluctuent d'une série statistique à l'autre.

Ces valeurs observées doivent donc être « compensées » afin de devenir les « vraies valeurs », c'est-à-dire les valeurs les plus probables. En conséquence, on tentera toujours d'identifier légitimement, hypothèses et expériences à l'appui, la distribution observée à l'une des distributions caractéristiques étudiées de façon générale dans la théorie de la statistique mathématique.

Il ne convient pas ici de se livrer à une étude détaillée des multiples méthodes conduisant à une solution convenable du problème. Disons seulement qu'elles conduisent toutes à des solutions approchées, solutions dont il est possible de pousser, dans bien des cas, l'approximation aussi loin qu'on le désire.

Notons dès maintenant que, dans le choix des approximations, une juste mesure s'impose. Si une approximation grossière n'est pas de mise, une approxi-

mation exagérée devient illusoire, car elle revient à apporter des corrections analytiques insignifiantes par rapport aux écarts entre les grandeurs issues de la mesure directe et les grandeurs théoriques.

§ 3. Moyennes et écarts. Moments.

Il sera exposé ici, sommairement, les méthodes essentielles susceptibles d'application dans cette étude biométrique et dans les études similaires. Toute démonstration mathématique a été exclue et il est demandé au lecteur d'accepter les résultats énoncés. Néanmoins, il a paru indispensable de définir un certain nombre de notions d'usage très fréquent.

1. Distributions discontinues :

Utilisant les notations déjà rencontrées au § 1, nous définirons la moyenne arithmétique des x_i par

$$A = \frac{1}{N} \sum_{i=1}^n x_i \varphi(x_i).$$

La quantité $x_i - A$ est l'écart absolu de la moyenne arithmétique.

La dispersion ou écart quadratique moyen σ est la racine carrée de la moyenne arithmétique des carrés des écarts absolus :

$$\sigma^2 = \frac{1}{N} \sum_{i=1}^n (x_i - A)^2 \varphi(x_i).$$

Le moment d'ordre s de la variable x est, par définition, la moyenne arithmétique des $s^{\text{ièmes}}$ puissances des x_i :

$$M_s = \frac{1}{N} \sum_{i=1}^n x_i^s \varphi(x_i).$$

Le moment d'ordre s de l'écart absolu est, de même,

$$\mu_s = \frac{1}{N} \sum_{i=1}^n (x_i - A)^s \varphi(x_i).$$

2. Distributions continues :

Les mêmes définitions se retrouvent dans le cas des distributions continues, mais les signes d'intégration \int sont à substituer aux signes de sommation Σ .

La moyenne arithmétique ou valeur moyenne est

$$A = \frac{1}{N} \int_a^b x p(x) dx = \int_a^b x T_x dx$$

et l'on retrouve sans peine l'analogie avec la définition correspondante du 1. En effet, $p(x)dx$ est le nombre des valeurs de x comprises dans un intervalle (infini-

ment petit) dx , centré sur le point x ; $x p(x) dx$ est la somme de ces valeurs de x , aux infiniment petits du second ordre près, et

$$\int_a^b x p(x) dx$$

est la somme de ces valeurs étendue à tout l'intervalle (ab) .

Enfin,

$$\frac{1}{N} \int_a^b x p(x) dx$$

est la moyenne de ces N valeurs.

De la même façon, l'écart quadratique moyen est

$$\sigma^2 = \frac{1}{N} \int_a^b (x - A)^2 p(x) dx = \int_a^b (x - A)^2 T_x dx.$$

Les moments M_s et μ_s de la variable x et de l'écart $x-A$ sont respectivement

$$M_s = \frac{1}{N} \int_a^b x^s p(x) dx = \int_a^b x^s T_x dx$$

et

$$\mu_s = \frac{1}{N} \int_a^b (x - A)^s p(x) dx = \int_a^b (x - A)^s T_x dx.$$

En particulier, qu'il s'agisse d'une distribution continue ou d'une distribution discontinue, on a

$$\begin{aligned} M_0 &= 1; & M_1 &= A \\ \mu_0 &= M_0 = 1; & \mu_1 &= 0; & \mu_2 &= \sigma^2. \end{aligned}$$

3. Correction des moments :

Dans le cas d'une distribution continue, on ne connaît pas, à priori, la forme de la fonction T_x . Tout au plus peut-on se permettre certaines hypothèses au sujet de cette fonction, hypothèses vraisemblables en raison de confirmations très fréquentes. On peut calculer des moments « observés » m_s par la formule

$$m_s = \frac{1}{N} \sum_{i=1}^n x_i^s \varphi(x_i) h,$$

et les hypothèses légitimes sur la fonction T_x permettent d'exprimer les moments théoriques M_s en fonction des moments observés. On est conduit ainsi aux formules de SHEPPARD [R. RISSER et C. E. TRAYNARD ⁽⁵³⁾], d'emploi fréquent. On a

$$\begin{aligned} M_0 &= m_0; & M_1 &= m_1; \\ M_2 &= m_2 - \frac{1}{12} m_0 h^2; \\ M_3 &= m_3 - \frac{1}{4} m_1 h^2, \text{ etc.} \end{aligned}$$

⁽⁵³⁾ RISSER, R. et TRAYNARD, C. E., 1933, pp. 35-37.

On ne tiendra compte des termes de correction que pour autant qu'ils ne soient pas négligeables vis-à-vis des moments observés.

4. Signification de la dispersion :

Dans les deux cas de distributions continues et discontinues, considérons les expressions

$$S = \int_a^b (x - X)^2 T_x dx \text{ et } S = \frac{1}{N} \sum_{i=1}^n (x_i - X)^2 \varphi(x_i),$$

où X est une quantité variable. Les valeurs de X rendant minimum ces expressions obéissent dans les deux cas à l'équation $\frac{dS}{dX} = 0$, ou bien, en divisant au préalable par -2 , à

$$\int_a^b (x - X) T_x dx = 0$$

ou

$$\frac{1}{N} \sum_{i=1}^n (x_i - X) = 0,$$

ou bien encore

$$X = \int_a^b x T_x dx = A,$$

$$X = \frac{1}{N} \sum_{i=1}^n x_i = A.$$

La moyenne des carrés des écarts à la moyenne arithmétique est plus petite que la moyenne des carrés des écarts à tout autre nombre.

De plus la forme même de σ^2 indique, et un raisonnement rigoureux du calcul des probabilités le démontre, que plus σ est petit, plus la répartition est concentrée autour de la valeur moyenne, d'où le nom de dispersion.

Il convient de considérer, non pas tant la valeur absolue de σ que sa valeur relative par rapport aux x_i , ou par rapport au nombre A qui en caractérise l'ordre de grandeur. A cet effet, on utilise souvent, comme mesure de la concentration, le coefficient de concentration $\frac{\sigma}{A}$. On suppose les x_i essentiellement positifs, sans quoi on pourrait obtenir $A=0$ et ce coefficient serait dépourvu de sens.

5. Moments d'ordre supérieur :

Ils jouent un rôle de première importance dans les diverses solutions des problèmes posés par l'ajustement des séries statistiques à une variable. Les fonctions qui représentent théoriquement les distributions observées sont, en effet, très souvent exprimées au moyen de sommes de fonctions simples dont les coefficients s'expriment au moyen de ces moments. L'exposé, très succinct, du § 4 en donne une explication.

§ 4. Principe des moments.

Il s'agit d'un principe sur lequel on peut fonder la solution du problème de l'ajustement. Il revient à déterminer une fonction $T_x^{(n)}$, telle que ses $n+1$ premiers moments des écarts $\mu_0, \mu_1, \dots, \mu_n$ soient égaux à ceux des écarts observés (éventuellement corrigés au moyen des formules de SHEPPARD).

On devra donc réaliser les conditions

$$\mu_s = \int_{-\infty}^{+\infty} x_s T_x^{(n)} dx = \int_{-\infty}^{+\infty} x_s T_x dx; \quad s = 0, 1, 2, \dots, n,$$

où T_x représente le taux théorique, $T_x^{(n)}$ le taux obtenu par l'application du principe des moments. $T_x^{(n)}$ constitue donc une approximation de T_x et, de façon générale, cette approximation est d'autant meilleure que le nombre $n+1$ de moments est grand. Cette approximation est, suivant une première méthode, obtenue au moyen de développements en série limités à un nombre fini de termes.

Cette première méthode est celle du développement en série de polynômes d'HERMITE [R. RISSER et C. E. TRAYNARD ⁽⁵⁴⁾, R. DELTHEIL ⁽⁵⁵⁾]. Sous certaines conditions de continuité et de dérivabilité, toute fonction $y=f(\xi)$ peut s'écrire sous forme d'un développement convergent :

$$y = (a_0 + a_1 H_1 + a_2 H_2 + \dots) e^{-k^2 \xi^2},$$

où

$$H_n = e^{+k^2 \xi^2} \frac{d^n}{d\xi^n} (e^{-k^2 \xi^2})$$

est un polynôme de degré n .

On a

$$H_0 = 1, H_1 = 2 k^2 \xi, H_2 = 4 k^4 \xi^2 - 2 k^2, \dots \text{ etc.}$$

Les propriétés analytiques de ces fonctions H_n permettent de démontrer que le développement dans l'intervalle $(-\infty, +\infty)$

$$y = \frac{1}{\sqrt{2\pi\mu_2}} e^{-\frac{\xi^2}{2\mu_2}} \left[1 - \frac{\mu_3}{6} H_3 + \left(\frac{\mu_4}{24} - \frac{\mu_2}{8} \right) H_4 + \dots \right] \quad (1)$$

convient à de nombreuses lois de probabilité $y=T_x$. Il suffit de poser $\xi=x-A$ et d'y remplacer les μ_s par les moments des écarts, tirés de l'observation.

Une seconde méthode, très précise et conduisant à d'excellents résultats, est celle de PEARSON. Elle est basée en premier lieu sur une propriété établie par

⁽⁵⁴⁾ RISSER, R. et TRAYNARD, C. E., 1933, pp. 218-221.

⁽⁵⁵⁾ DELTHEIL, R., 1930, pp. 87-90.

PEARSON, revenant à dire que de nombreuses lois de probabilité $y=T_x$ obéissent à l'équation différentielle

$$\frac{dy}{y} = \frac{x-a}{F(x)} dx, \quad (2)$$

où $F(x)$ est un polynôme de la forme

$$F(x) = \alpha_0 + \alpha_1 x + \alpha_2 x^2 + \dots$$

On a alors, en général,

$$y = C e^{\int \frac{x-a}{F(x)} dx},$$

C étant déterminé par le fait que l'intégrale définie de y dans le domaine de x est égale à 1 (relation de certitude mathématique).

On peut démontrer que la connaissance des moments observés permet de déterminer $a, \alpha_0, \alpha_1, \alpha_2, \dots$, etc. En particulier, si l'on connaît M_1 et M_2 , on peut déterminer a et α_0 dans l'équation

$$\frac{dy}{y} = \frac{x-a}{\alpha_0} dx. \quad (3)$$

On trouve, en effet,

$$a = M_1, \alpha_0 = M_1^2 - M_2$$

et

$$y = \frac{1}{\sqrt{2\pi(M_2 - M_1^2)}} e^{-\frac{(x - M_1)^2}{2(M_2 - M_1^2)}}. \quad (4)$$

C'est la solution du problème dit des deux moments.

L'utilisation de trois, quatre... moments conduit à des solutions plus générales et plus précises.

On obtient alors des lois ou taux de probabilité dont la représentation graphique conduit aux courbes de PEARSON [R. RISSER et C. E. TRAYNARD ⁽⁵⁶⁾].

§ 5. La loi de GAUSS.

L'application de l'une des méthodes précédentes à la solution de problèmes d'ajustement statistique permet donc d'obtenir des solutions de forme très générale, et la précision de la représentation analytique peut être poussée aussi loin qu'on le désire, si l'on veut pousser le calcul des moments d'ordre supérieur jusqu'à un ordre suffisamment élevé. Notons toutefois que cette précision devient tout à fait illusoire à partir de l'instant où la correction apportée par le calcul de moments d'ordres plus élevés est inférieure à l'erreur de mesure et, à fortiori, à l'erreur arithmétique commise en négligeant des termes à partir d'un certain ordre de grandeur très petit.

⁽⁵⁶⁾ RISSER, R. et TRAYNARD, C. E., 1933, pp. 52-62.

D'autre part, on admet couramment, à titre de postulat, que les écarts des valeurs mesurées à la « vraie valeur » de la grandeur mesurée sont distribués symétriquement autour de cette « vraie valeur ». Autrement dit, on admet des probabilités égales pour des écarts de signes contraires, mais égaux en valeur absolue. Enfin, si l'on pose que la « vraie valeur » est la moyenne arithmétique des valeurs observées, on justifie sans peine la loi de distribution bien connue sous le nom de loi de GAUSS.

Cette loi n'est pas une loi générale de distribution des grandeurs mesurées autour d'une valeur remarquable, dénommée « vraie valeur », aussi remarquable que dépourvue d'une définition précise. Elle est le résultat d'une justification excellente, mais non d'une démonstration, basée sur des hypothèses ayant la valeur de postulats, qu'on est libre, en toute logique, de rejeter et de remplacer par d'autres, plus complexes.

L'expression mathématique de la loi de GAUSS est la suivante :

$$y = T_x = \frac{1}{\sigma \sqrt{2\pi}} e^{-\frac{(x-A)^2}{2\sigma^2}},$$

ou encore

$$p(x) = \frac{N}{\sigma \sqrt{2\pi}} e^{-\frac{(x-A)^2}{2\sigma^2}}. \quad (5)$$

Les hypothèses sur lesquelles cette loi est fondée, si elles ne manquent pas de légitimité, manquent, comme on le voit, de généralité dans bien des cas d'application.

Ce manque de généralité se remarque d'ailleurs aussi dans les exposés sommaires des méthodes des polynômes d'HERMITE et de PEARSON, rappelées précédemment.

En effet, dans le développement de la formule (1), si l'on néglige les coefficients de H_3 et H_4 , on retrouve

$$y = \frac{1}{\sqrt{2\pi\mu_2}} e^{-\frac{(x-A)^2}{2\mu_2}},$$

où

$$\sigma^2 = \mu_2$$

et la solution de l'équation de PEARSON (3), cas particulier de la formule (2), est également l'expression de la loi de GAUSS.

On a, en effet,

$$\begin{aligned} \sigma^2 = \mu_2 &= \int_a^b (x-A)^2 T_x dx \\ &= \int_a^b x^2 T_x dx - 2A \int_a^b x T_x dx + A^2 \\ &= M_2 - 2A^2 + A^2 = M_2 - A^2 = M_2 - M_1^2. \end{aligned}$$

Il convient d'insister fortement sur ce fait : la loi de GAUSS n'est qu'un cas très particulier des lois de distribution des séries statistiques à une variable. Elle n'a pas le caractère universel de loi des erreurs de mesure qu'une habitude et une tradition de l'emploi lui confèrent, d'une façon implicite et quasi incontestable, dans l'esprit de nombreuses personnes.

Si son emploi est abondamment répandu, c'est que, dans nombre de cas, elle donne une bonne représentation de la réalité observée, sans qu'il soit toutefois permis de prétendre que ce soit la meilleure. La loi de GAUSS possède quelques propriétés analytiques qu'il est, en général, très aisé de confronter avec des propriétés correspondantes de la distribution observée. Si des écarts trop considérables sont constatés entre la théorie et l'expérience, il faut délibérément adopter une autre loi de distribution.

Dans l'étude de la distribution des indices caractéristiques des populations d'*Anthraconauta* faisant l'objet de ce travail et exposés au chapitre suivant, il a été fait un usage général de la loi de GAUSS, parce que l'ajustement pouvait être considéré comme satisfaisant et aussi parce que le nombre réduit des exemplaires étudiés (plusieurs centaines) ne permettait pas de tirer des conclusions tendant à rejeter cette loi.

Il nous est apparu cependant que les distributions observées étaient plus concentrées que la distribution de GAUSS. Nous pensons qu'une étude portant sur un très grand nombre de coquilles révélerait une distribution quasi symétrique, à concentration plus forte, ce qui s'exprimerait aisément par l'emploi de moments jusqu'au 4^e ordre dans les méthodes de développement en séries de polynômes d'HERMITE et de PEARSON.

§ 6. Propriétés de la loi de GAUSS.

1° De façon générale, nous avons vu que

$$\int_a^b p(x) dx = N \quad \text{ou} \quad \int_a^b T_x dx = 1,$$

l'intervalle (a, b) étant celui dans lequel se trouvent certainement toutes les valeurs de x .

Ici on a

$$\int_{-\infty}^{+\infty} \frac{1}{\sigma \sqrt{2\pi}} e^{-\frac{(x-A)^2}{2\sigma^2}} dx = 1,$$

ce qui, géométriquement, signifie que l'aire située sous la « courbe en cloche » de GAUSS, entre cette courbe et l'axe des x , de $-\infty$ à $+\infty$, est égale à 1. Les bornes de l'intervalle où se trouvent certainement toutes les valeurs de x sont $-\infty$ et $+\infty$. Il est bien évident qu'en pratique, les bornes de cet intervalle sont finies, mais on remarque qu'à partir d'un certain écart de la moyenne A , l'aire située sous la courbe en dehors de cet écart est pratiquement nulle.

La probabilité d'obtenir une valeur de x comprise entre x_1 et x_2 est

$$\int_{x_1}^{x_2} \frac{1}{\sigma \sqrt{2\pi}} e^{-\frac{(x-A)^2}{2\sigma^2}} dx.$$

Pour calculer ces probabilités, on opère la substitution $t = \frac{x-A}{\sigma \sqrt{2}}$.

De cette façon, la probabilité précédente s'écrit

$$\frac{1}{\sqrt{\pi}} \int_{t_1}^{t_2} e^{-t^2} dt. \quad (6)$$

On a dressé des tables donnant les valeurs de la fonction de Poisson :

$$\Theta(\lambda) = \frac{1}{\sqrt{\pi}} \int_{-\lambda}^{+\lambda} e^{-t^2} dt = \frac{2}{\sqrt{\pi}} \int_0^{\lambda} e^{-t^2} dt,$$

qui est la probabilité pour que

$$A - \lambda\sigma \sqrt{2} < x < A + \lambda\sigma \sqrt{2}.$$

Ces tables, établies pour un grand nombre de valeurs de λ , permettent le calcul des probabilités (6).

On constate, notamment, que pour

$$\begin{aligned} \lambda = 1, & \quad \Theta(\lambda) = 0,8427, \\ = 2, & \quad \Theta(\lambda) = 0,9953, \\ = 3, & \quad \Theta(\lambda) = 0,99998, \\ = 4, & \quad \Theta(\lambda) = 0,99999998. \end{aligned}$$

On peut admettre que, pratiquement, les écarts de la moyenne arithmétique ne dépassent pas en valeur absolue $3\sigma \sqrt{2}$, ou bien qu'on a, avec une quasi-certitude,

$$A - 3\sqrt{2}\sigma < x < A + 3\sqrt{2}\sigma.$$

2° Définissons à présent l'écart probable et l'écart moyen, éléments caractéristiques de la loi de GAUSS.

L'écart probable E_p est celui pour lequel les probabilités d'être dépassé et de ne pas être atteint sont égales, donc, toutes deux égales à $1/2$. Autrement dit, la moitié des valeurs observées, supposées obéir à la loi de GAUSS, se trouvent dans l'intervalle $(A - E_p, A + E_p)$.

Ayant

$$\frac{1}{2} = \frac{1}{\sqrt{\pi}} \int_{-\lambda}^{+\lambda} e^{-t^2} dt, \text{ pour } \lambda = 0,4769$$

on a

$$E_p = 0,4769 \sqrt{2} \sigma = 0,6745\sigma. \quad (7)$$

L'écart moyen E_m est la moyenne des valeurs absolues des écarts $x - A$:

$$E_m = \int_{-\infty}^{+\infty} \frac{1}{\sqrt{\pi}} |x - A| e^{-\frac{(x - A)^2}{2\sigma^2}} dx;$$

$$E_m = \sigma \frac{\sqrt{2}}{\pi} = 0,7977\sigma. \quad (8)$$

Les constantes σ_x et E_p sont des mesures commodes pour les écarts de la loi de GAUSS.

D'autre part, les relations précédentes entre σ_x , E_m et E_p sont des moyens de contrôle du bien-fondé de l'identification d'une répartition expérimentale à cette loi de distribution.

§ 7. De la moyenne arithmétique considérée comme une variable statistique.

Soit une série de N expériences portant sur une variable statistique x et ayant fourni les valeurs x_1, x_2, \dots, x_n , certaines valeurs pouvant se reproduire plusieurs fois.

Posons $z = \frac{x_1 + x_2 + \dots + x_n}{N}$; chaque série de N épreuves sur le compte de x fournit des valeurs x_1, x_2, \dots, x_n et une certaine valeur de z , z est une variable statistique au même titre que les x_1, x_2, \dots, x_n , et nous l'appellerons la variable moyenne.

Affectant des indices x et z les symboles des moyennes et dispersions, il est aisé de démontrer que

$$A(z) = A(x), \quad (9)$$

$$\sigma_z^2 = \frac{\sigma_x^2}{N},$$

ou bien

$$\sigma_z = \frac{\sigma_x}{\sqrt{N}}. \quad (10)$$

Si l'on connaît le taux T_x , on peut calculer T_z et, notamment, on peut démontrer que si la variable x obéit à la loi de GAUSS, il en est de même de la variable moyenne z . On obtient alors, de façon précise,

$$T_z = \frac{1}{\sigma_z \sqrt{2\pi}} e^{-\frac{(z - A)^2}{2\sigma_z^2}}. \quad (11)$$

Si x ne suit pas rigoureusement cette loi de distribution, T_z n'y obéit pas, mais la formule [10] montre que si N croît indéfiniment, σ_z a pour limite 0. Ceci suffit pour prouver que si le nombre d'épreuves croît indéfiniment, la

variable moyenne z est « concentrée » en $A(x)$, ce qui signifie que l'ensemble des valeurs quasi certaines de z est constitué par un intervalle tendant vers 0 et centré sur $A(x)$.

Dans ces conditions, que nous qualifierons d'asymptotiques, il est très approché d'admettre que z obéit à la loi de GAUSS [11].

Dans le cas de la présente étude, cette approximation est excellente et d'autant meilleure que N est plus grand.

On peut alors admettre que l'intervalle des valeurs quasi certaines de z est défini par

$$A(x) - 3 \sqrt{2} \frac{\sigma_x}{\sqrt{N}} < Z < A(x) + 3 \sqrt{2} \frac{\sigma_x}{\sqrt{N}} \quad (12)$$

et que

$$3 \sqrt{2} \frac{\sigma_x}{\sqrt{N}} = E \quad (13)$$

constitue, en valeur absolue, une limite supérieure de l'erreur commise sur la moyenne.

Cette conclusion est d'une grande importance pratique, comme nous aurons l'occasion de le voir.

Comme au § 6, nous obtiendrons les expressions de l'écart probable sur la moyenne $E_p^{(m)}$:

$$E_p^{(m)} = 0,6745 \frac{\sigma_x}{\sqrt{N}}, \quad (14)$$

et de l'écart moyen sur la moyenne

$$E_m^{(m)} = 0,7977 \frac{\sigma_x}{\sqrt{N}}. \quad (15)$$

§ 8. Symétrie et asymétrie d'une distribution observée.

La notion intuitive de la symétrie d'une distribution est très claire : une distribution est symétrique si les valeurs de la variable étudiée se répartissent symétriquement autour d'une valeur particulière.

Analytiquement, la distribution de la variable x est symétrique autour de a si la fonction T_x est paire en $x - a$:

$$T_x = f[(x - a)^2],$$

c'est-à-dire si T_x ne s'exprime qu'au moyen des puissances paires de $x - a$, et réciproquement. On démontre alors que $A = a$.

Géométriquement, la courbe $y = T_x$ est symétrique par rapport à l'axe $x = A$ (fig. 4).

Une propriété remarquable s'en dégage : tous les moments des écarts d'ordre impair d'une distribution symétrique sont nuls :

$$\mu_1 = \mu_3 = \mu_5 = \dots \mu_{2n+1} = 0.$$

La loi de GAUSS, en particulier, est une distribution symétrique. Il est aisé d'obtenir une forme analytique générale de toutes les lois de distributions symétriques dans l'intervalle de $+\infty$ à $-\infty$. Il suffit, dans le développement en série de polynômes d'HERMITE, d'annuler tous les termes contenant $\mu_1, \mu_3, \mu_5, \dots, \mu_{2n+1}$.

En pratique, il est rare, parce que très peu probable, d'obtenir une répartition observée rigoureusement symétrique. En effet, même si l'on sait, à priori, que la répartition est théoriquement symétrique, ou si l'on a des raisons légi-

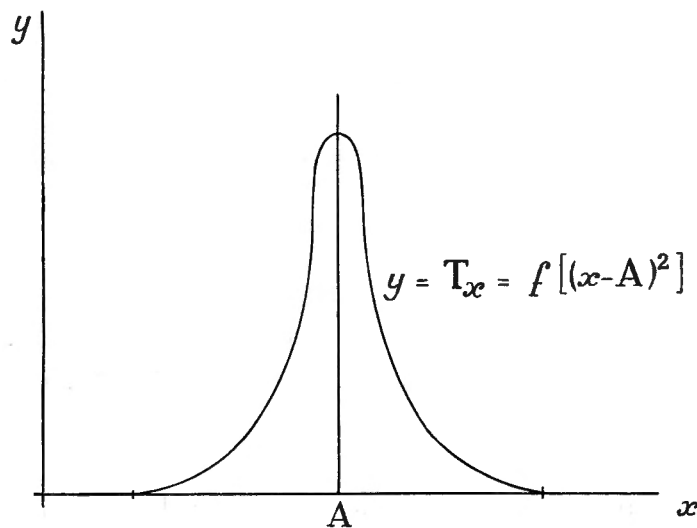


FIG. 4.

times de le supposer, la loi des écarts dus au hasard ou aux erreurs joue toujours lorsqu'on tente une réalisation expérimentale. On ne peut donc affirmer l'asymétrie d'une distribution sur la foi de l'observation montrant une asymétrie modérée. En fait, l'hypothèse légitime décide de la symétrie et l'hypothèse ne peut être infirmée que par des divergences flagrantes.

La meilleure mesure pratique est la constatation de la non-nullité des moments μ_{2n+1} , ou, du moins, du fait que ceux-ci sont, en valeur absolue, d'ordres de grandeurs comparables aux moments d'ordre pair.

Les mesures empiriques de l'asymétrie (« skewness » des Anglo-Saxons) ont un caractère synthétique; mais, comme d'autres constantes statistiques, elles ne donnent qu'une indication générale imprécise. On peut trouver un exposé détaillé de la méthode de ces mesures dans le traité remarquable de G. UDNY YULE et M. G. KENDALL ⁽⁵⁷⁾.

⁽⁵⁷⁾ UDNY YULE, G. et KENDALL, M. G., 1945, pp. 96, 98, 162-164.

§ 9. Méthode pratique de calcul des moyennes et de la dispersion.

Soit une variable statistique continue x susceptible de prendre toutes les valeurs d'un certain intervalle. Supposons que celui-ci soit, pour les nécessités du relevé des fréquences, partagé en intervalles partiels de longueur h .

Comment fixer le choix de h ? Si l'on désire se rapprocher des exigences théoriques, il faudrait que h fût infiniment petit. D'autre part, il faut aussi que l'intervalle considéré contienne des valeurs de x donnant lieu à des observations, car le nombre des valeurs observées de x est forcément fini. On prendra donc un nombre d'intervalles dépendant de l'intervalle total, ni trop petit, pour que la courbe à paliers garde une signification, ni trop grand, pour qu'on puisse effectivement conserver des valeurs de x et pour donner plus de régularité à ce qu'on observe. Il faut noter que ces intervalles doivent, au moins, être d'une grandeur double de l'erreur de mesure.

En général on prendra de dix à vingt intervalles.

Les calculs de moyennes, de dispersions et, en général, de moments peuvent être très longs, si l'on emploie simplement les formules du § 3. Aussi opère-t-on souvent une substitution de variable très simple, ramenant à l'unité l'intervalle de longueur h . On passe de la variable x à une variable X par la relation

$$x = hX + x_0, \quad (16)$$

où x_0 est une valeur de x convenablement choisie, voisine, à première vue, de la moyenne arithmétique, de sorte que X prendra des valeurs positives et négatives, assez petites, dont la moyenne sera voisine de 0.

On voit que si $x_2 - x_1 = h$, on a $h(X_2 - X_1) = h$ et $X_2 - X_1 = 1$.

L'intervalle h est donc bien ramené à l'unité.

Cela étant, désignant par $A(x)$, σ_x , $A(X)$, σ_X les moyennes et écarts quadratiques de x et X , on a immédiatement

$$A(x) = h A(X) + x_0. \quad (17)$$

Soustrayant [17] de [16], on a $x - A(x) = h[X - A(X)]$ et $\sigma_x^2 = h^2 \sigma_X^2$, ce qui ne dépend pas de x_0 .

Ces formules sont d'un usage facile et nous en ferons un emploi fréquent.

Par exemple, soit une variable statistique x observée dans l'intervalle $113,5 < x < 157,5$, soit $h = 4$, d'où le nombre d'intervalles partiels $n = 11$.

On a $x_1 = 113,5 + \frac{h}{2} = 115,5$, $x_2 = 119,5$, $x_3 = 123,5$; ... $x_{10} = 151,5$, $x_{11} = 155,5$.

Posons $x = 4X + 135,5$. On trouve alors

$X_1 = -5$	$X_5 = -1$	$X_9 = 3$
$X_2 = -4$	$X_6 = 0$	$X_{10} = 4$
$X_3 = -3$	$X_7 = 1$	$X_{11} = 5$
$X_4 = -2$	$X_8 = 2$	

Soit, enfin, le relevé des fréquences :

$\varphi(x_1)h = 2$	$\varphi(x_5)h = 37$	$\varphi(x_9)h = 5$
$\varphi(x_2)h = 2$	$\varphi(x_6)h = 45$	$\varphi(x_{10})h = 1$
$\varphi(x_3)h = 7$	$\varphi(x_7)h = 27$	$\varphi(x_{11})h = 1$
$\varphi(x_4)h = 14$	$\varphi(x_8)h = 18$	

$$\sum_{i=1}^{11} \varphi(x_i)h = 159.$$

On a

$$A(x) = \frac{1}{159} [2 \times (-5) + 2 \times (-4) + 7 \times (-3) + 14 \times (-2) + 37 \times (-1) + 45 \times 0 + 27 \times 1 + 18 \times 2 + 5 \times 3 + 1 \times 4 + 1 \times 5] = \frac{17}{159} = 0,106$$

et

$$A(x) = 4X(-0,1) + 135,5 = 135,1.$$

Calculer les carrés $X_i - A(X)$ est fastidieux. On dispose d'une autre relation simplifiant encore le calcul de la dispersion. Ayant en général

$$\sigma_x^2 = \int_a^b (x - A)^2 T_x dx,$$

ou bien

$$\sigma_x^2 = \int_a^b (x^2 - 2Ax + A^2) T_x dx,$$

il vient

$$\begin{aligned} \sigma_x^2 &= \int_a^b x^2 T_x dx - 2A \int_a^b x T_x dx + A^2 \int_a^b T_x dx. \\ \sigma_x^2 &= M_2 - 2A^2 + A^2, \\ \sigma_x^2 &= M_2 - A^2 = A(x^2) - [A(x)]^2. \end{aligned} \tag{19}$$

En particulier,

$$\begin{aligned} \sigma_x^2 &= A(X^2) - [A(X)]^2 = \frac{1}{159} [2 \times 25 + 2 \times 16 + 7 \times 9 + 14 \times 4 \\ &+ 37 \times 1 + 27 \times 1 + 18 \times 4 + 5 \times 9 + 1 \times 16 + 1 \times 25] - 0,0112 = 2,6484 \end{aligned}$$

et

$$\begin{aligned} \sigma_x &= 1,62 \\ \sigma_x &= 4 \times 1,62 = 6,48. \end{aligned}$$

§ 10. Liaisons entre deux variables statistiques. Corrélation.

Supposons que, dans un ensemble d'objets de même espèce, la mesure d'un caractère bien défini soit désignée par x , et celle d'un autre caractère par y , ces caractères étant supposés linéaires. Pour chaque spécimen on obtiendra ainsi un couple de grandeurs (x, y) et la question à résoudre, exprimée sous une forme

générale et toutefois imprécise, est la suivante : existe-t-il une liaison entre les caractères x et y mesurés sur chaque objet ?

Pour y répondre, certaines notions d'analyse mathématique et de calcul des probabilités doivent être rappelées et précisées.

Considérons, de façon générale, une collection d'objets ou d'individus d'une espèce, sur lesquels nous mesurons deux caractères, x et y .

Il peut se faire qu'une valeur x fixe y de manière univoque et réciproquement. On dit alors qu'il y a une liaison ou dépendance fonctionnelle entre x et y comme entre l'intensité du courant et la tension dans un circuit électrique de résistance invariable.

Il peut se faire aussi que cette dépendance soit plus ou moins lâche, en ce sens que pour chaque valeur de x ou chaque petit intervalle de valeurs de x on observe une répartition de valeurs de y ne permettant pas de conclure à une liaison fonctionnelle. Précisons cette notion par des exemples bien connus : Si l'on mesure la longueur et la largeur d'œufs de poule d'une espèce bien déterminée, le sous-groupe qu'on obtient en fixant la longueur, ou en considérant un très petit intervalle de longueurs, n'est pas composé seulement d'œufs ayant une largeur bien déterminée, ou comprise entre des limites très rapprochées, mais il y a aussi, pour chaque longueur, une répartition des largeurs suivant une loi de fréquence rappelant la loi de GAUSS. Il en est de même si l'on mesure la taille et la largeur d'épaules d'un groupe d'hommes assez nombreux. Le sous-groupe des hommes de même taille n'est pas composé d'hommes ayant, de surcroît, la même largeur d'épaules; mais on constate également une répartition d'allure symétrique rappelant la répartition suivant la loi de GAUSS.

On dit alors qu'il y a liaison stochastique entre les deux variables statistiques, longueur et largeur des œufs ou taille et largeur d'épaules, et d'une façon générale entre x et y .

Mais deux cas encore sont à distinguer : 1° la répartition des largeurs des œufs (ou des largeurs d'épaules) ne change pas, quelle que soit la longueur des œufs (ou les tailles des hommes); on dit alors que les deux variables x et y sont indépendantes au sens des probabilités; 2° la répartition des largeurs dépend de la longueur, et dans ce cas x et y , longueurs et largeurs, sont en corrélation.

Ces notions ne peuvent être exprimées avec précision que sous la forme analytique.

§ 11. Distribution des fréquences d'un complexe
de deux variables statistiques.

Il convient, pour avoir une idée plus exacte de la corrélation, de poser quelques définitions et de détailler l'origine des nombres par lesquels on tente d'en donner une mesure.

Considérons le cas de deux variables, x et y , continues, analogues aux longueurs et largeurs d'œufs ou aux tailles et largeurs d'épaules, en négligeant les cas où les deux variables sont discontinues, ou bien, où l'une est continue, l'autre discontinue. A chaque mesure, sur un même individu, des deux caractères x et y , faisons correspondre un point d'un plan rapporté aux axes O_x, O_y . Nous obtenons

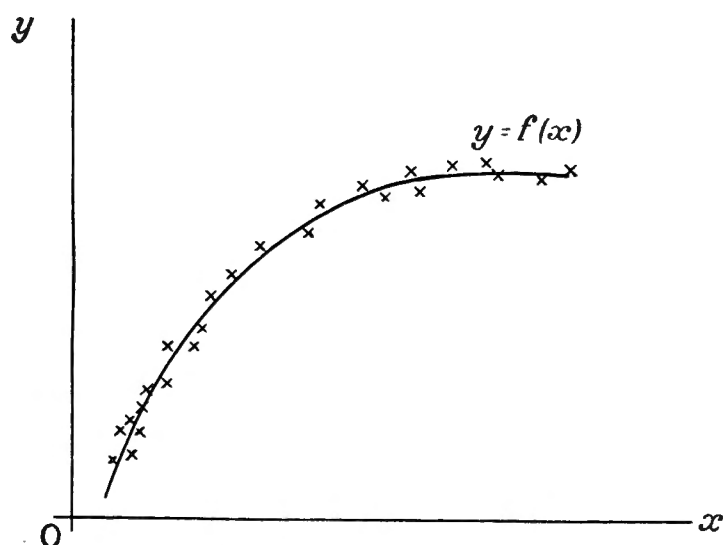


FIG. 5. — Liaison fonctionnelle.

ainsi un nuage de points dans le plan. Si la liaison est fonctionnelle, tous les points sont situés sur une même courbe ou au voisinage immédiat de celle-ci. Dans la négative, on peut observer une répartition des points avec une concentration plus ou moins prononcée autour d'une ligne. Si cette concentration est très prononcée, la liaison stochastique est proche de la liaison fonctionnelle (fig. 5 et fig. 6). Autrement, en découpant, dans le plan, une tranche de largeur suffisamment mince, perpendiculaire à O_x , on observe une certaine répartition de fréquence de y , variable en général, avec l'abscisse de la tranche choisie. C'est le cas de la figure 6. Si cette répartition ne varie pas, x et y sont indépendantes (fig. 7).

Traduisons analytiquement ces notions intuitives.

Supposons que les couples de valeurs observées (x, y) définissent des points situés dans un domaine D du plan xOy . Couvrons ce plan d'un quadrillage, et

soient h et k les côtés des rectangles parallèles aux axes O_x et O_y . Soient x_i, y_j les coordonnées du centre d'un quelconque des rectangles obtenus (fig. 8).

Désignons par $\psi(x_i, y_j)h.k$ le nombre des points du rectangle de centre (x_i, y_j)

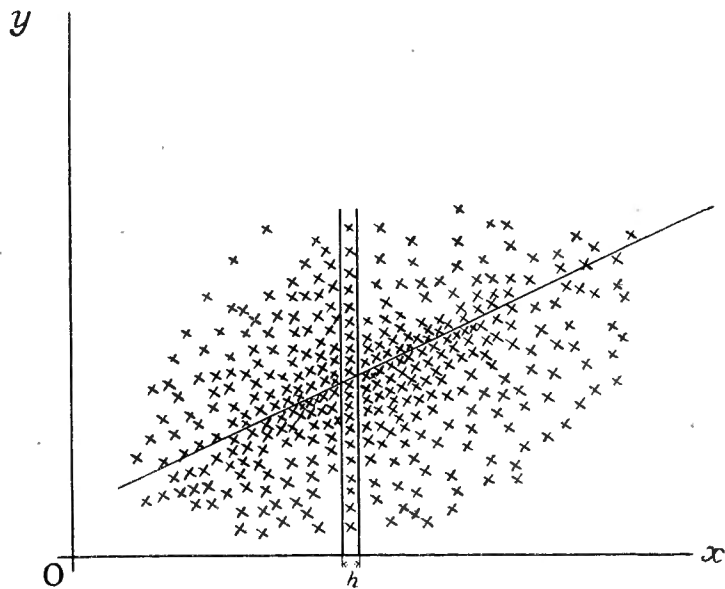


FIG. 6. — Corrélation.

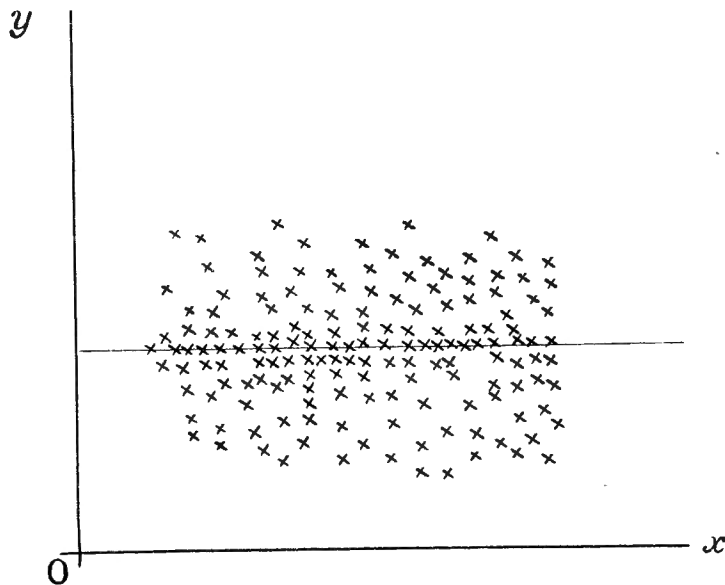


FIG. 7. — Indépendance au sens des probabilités.

effectivement dénombrés au cours du relevé. Si N individus ont fait l'objet de la mesure des deux caractères x et y , on a évidemment, les sommes Σ_i et Σ_j étant étendues à toutes les valeurs de i et j ,

$$\Sigma_i \Sigma_j \psi(x_i, y_j) h, k = N.$$

On définit ainsi une fonction discontinue $z = \psi(x, y)$, dont la représentation dans un système de trois axes rectangulaires est une surface constituée par la juxtaposition de parallélépipèdes droits, de hauteurs $\psi(x_i, y_j)$ et de bases égales, de côtés h, k (fig. 9).

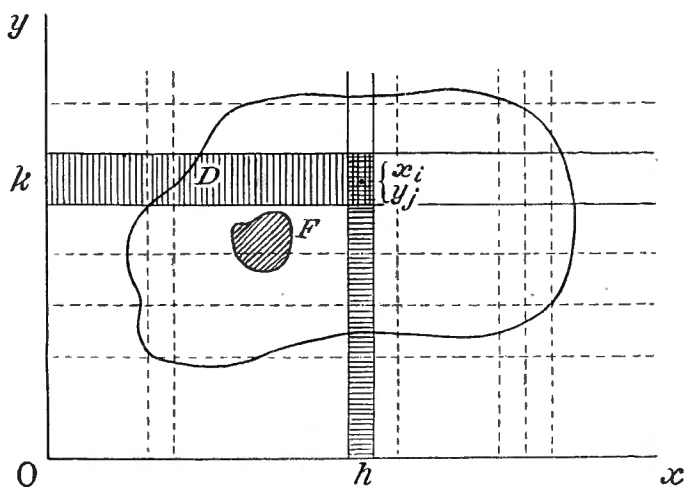


FIG. 8.

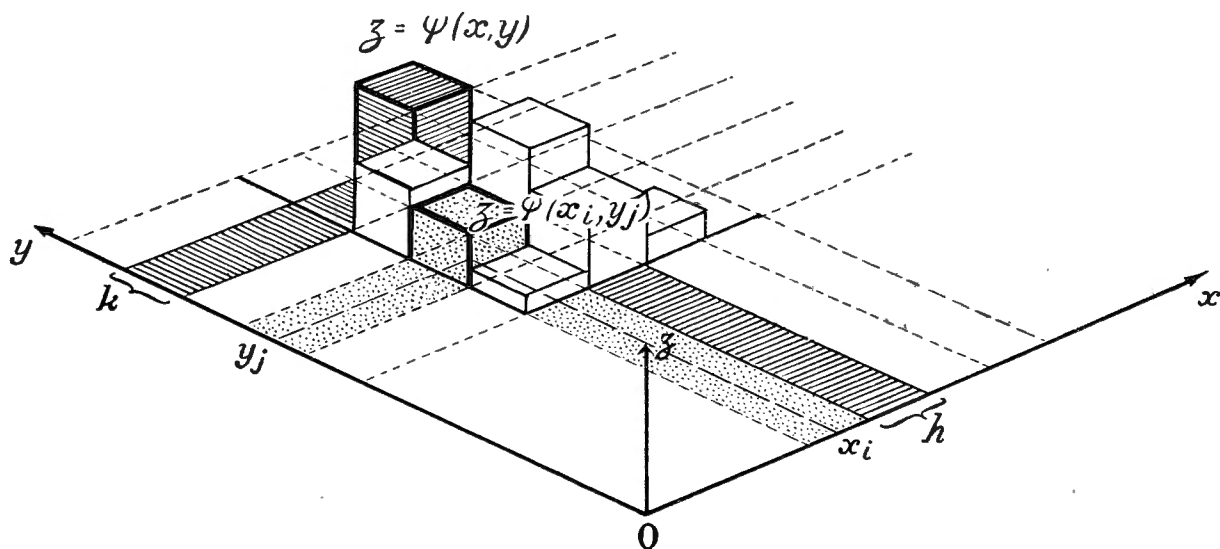


FIG. 9.

Ici également se pose, de façon générale, le problème de l'ajustement des distributions à deux variables. Il peut être résolu par des méthodes issues du calcul des probabilités, et sa solution revient à définir une fonction $p(x, y)$ ou une fonction $T_{xy} = \frac{p(x, y)}{N}$, en principe continue, telles que

$$\iint_D p(x, y) dx dy = N, \text{ ou } \iint_D T_{xy} dx dy = 1,$$

et que

$$\iint_F p(x, y) dx dy \text{ et } \iint_F T_{xy} dx dy$$

soient les nombres des points observés dans le domaine F, compris dans la limite de D, et la probabilité statistique d'obtenir un point compris dans les limites du domaine F (fig. 8).

La fonction T_{xy} est le taux de probabilité des deux variables x et y , dans le domaine D.

§ 12. Dépendance et indépendance au sens des probabilités.

Nous avons déjà vu, au § 1 et au § 2, de quelle façon peut se présenter la liaison entre les caractères x et y , et nous avons utilisé les termes « dépendance fonctionnelle » pour qualifier la liaison entre x et y , s'exprimant par une relation $y=f(x)$.

Il s'agit d'une dépendance s'exprimant sous la forme d'une fonction au sens de l'analyse mathématique.

La dépendance au sens de la théorie des probabilités s'exprime d'une façon différente [P. VAN DEUREN ⁽⁵⁸⁾].

Considérons deux événements, A et B, et un événement constitué par l'arrivée simultanée des deux événements A et B. Désignons par AB cet événement « composé ». Un principe fondamental du calcul des probabilités s'exprime par la relation

$$\begin{aligned} p_{AB} &= p_A p_{B:A} \\ &= p_B p_{A:B}, \end{aligned} \quad (19)$$

où p_{AB} est la probabilité d'arrivée de AB; p_A , p_B étant les probabilités d'arrivée de A et de B; $p_{A:B}$ la probabilité d'arrivée de A si B s'est produit, et $p_{B:A}$ la probabilité d'arrivée de B si A s'est produit.

Si A représente la constatation d'une valeur de x comprise dans l'intervalle $(x, x+dx)$, et B la constatation d'une valeur de y comprise dans l'intervalle $(y, y+dy)$, AB sera la constatation d'un couple (x, y) définissant un point intérieur au rectangle de côtés dx et dy , et l'on aura

$$\begin{aligned} T_{xy} &= T_x \cdot T_{y:x} \\ &= T_y \cdot T_{x:y}. \end{aligned} \quad (20)$$

Les taux de probabilités $T_{y:x}$ et $T_{x:y}$ sont les taux liés de y en x et de x en y .

$T_{y:x} dy$ est la probabilité d'obtenir une valeur de y entre y et $y+dy$, sachant que x est situé dans l'intervalle de x à $x+dx$. Autrement dit, $T_{y:x}$ est la loi de répartition des valeurs de y correspondant à la valeur x quelconque entre x et $x+dx$.

⁽⁵⁸⁾ VAN DEUREN, P., 1934, pp. 20, 73, 79-81.

En général, $T_{y:x}$ dépend de x et de y , c'est-à-dire que la loi de répartition de y en x varie avec x . On dit alors que x et y sont dépendantes au sens des probabilités.

Si, quelle que soit x , la loi de répartition de y en x reste invariable, autrement dit, si

$$T_{y:x} = T_y, \tag{21}$$

x et y sont indépendantes au sens des probabilités.

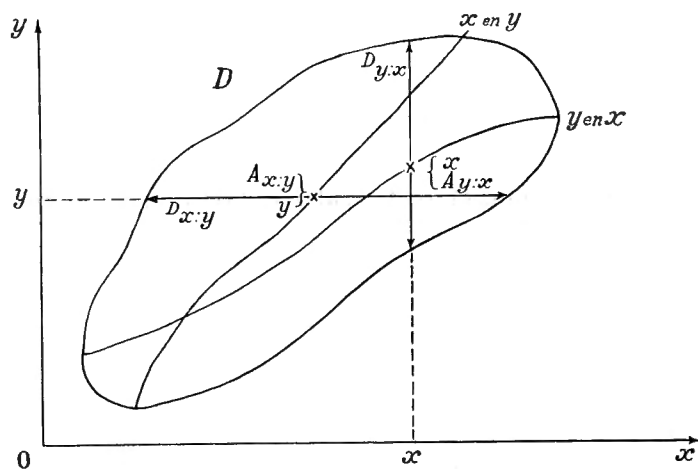


FIG. 10.

Notons que la relation [21], en vertu de [20], entraîne

$$T_{x:y} = T_x.$$

La dépendance fonctionnelle, analytique, et la corrélation entrent toutes deux dans la définition de la dépendance au sens de la théorie des probabilités.

§ 13. Moyennes liées. Lignes de régression.

Une valeur de x étant fixée, l'ensemble des valeurs de y , observées en même temps que cette valeur de x , est contenu dans un certain intervalle $D_{y:x}$ dont les extrémités dépendent de x , comme le montre intuitivement la figure 10, et l'on a

$$\int_{D_{y:x}} T_{y:x} dy = 1,$$

relation de certitude, exprimant que toutes les valeurs de y , x étant fixée, sont dans l'intervalle $D_{y:x}$.

Les moments de la loi de répartition $T_{y:x}$ sont les moments liés de y en x et, parmi eux,

$$A_{y:x} = \int_{D_{y:x}} y T_{y:x} dy \tag{22}$$

est la moyenne liée de y en x .

A chaque valeur de x correspond ainsi une moyenne liée de y en x , et le lieu géométrique des points de coordonnées $(x, A_{y:x})$ est la ligne de régression de y en x .

Cette ligne est telle que la somme des carrés $(y - A_{y:x})^2$, pour toutes les valeurs de x et y , est minimum.

On définirait de même la moyenne liée et la ligne de régression de x en y par

$$A_{x:y} = \int_{D_{x:y}} x T_{x:y} dx. \quad (23)$$

Les expressions [22] et [23], qui sont des fonctions de x et y respectivement, sont les équations des lignes de régression de y en x et de x en y .

§ 14. Ecart quadratique et rectangulaire. Coefficient de corrélation.

Si l'on ne considère que la seule variable x , indépendamment des valeurs correspondantes de y , on peut étudier la loi de distribution de x , ses moyennes et moments, et l'on démontre notamment que

$$A_x = \iint_D x T_{xy} dx dy \quad (24)$$

et, de même, que l'écart quadratique de x est défini par

$$\sigma_x^2 = \iint_D (x - A_x)^2 T_{xy} dx dy.$$

En fait, si l'on se réfère aux définitions données au § 3, on constate, sans entrer dans la voie des démonstrations analytiques, que A_x , calculée selon [24], est bien la moyenne de toutes les valeurs de x , la double intégration étant étendue à toutes les valeurs de x du domaine D .

En pratique, nous aurons à calculer A_x et σ_x , alors que nous sommes dans l'ignorance, à priori, de la forme de T_{xy} . Substituant à T_{xy} la fonction discontinue $\psi(x_i, y_j)$ définie par les valeurs (observées) qu'elle prend dans les rectangles de centres (x_i, y_j) et de côtés h et k , nous aurons

$$A_x = \frac{1}{N} \sum_i \sum_j x_i \psi(x_i, y_j) h \cdot k, \quad (25)$$

$$\sigma_x^2 = \frac{1}{N} \sum_i \sum_j (x_i - A_x)^2 \psi(x_i, y_j) h \cdot k. \quad (26)$$

En posant

$$\varphi(x_i) = \sum_j \psi(x_i, y_j) k,$$

on a

$$A_x = \frac{1}{N} \sum_i x_i \varphi(x_i) h, \quad (27)$$

$$\sigma_x^2 = \frac{1}{N} \sum_i (x_i - A_x)^2 \varphi(x_i) h. \quad (28)$$

$\varphi(x_i)h$ est le nombre de valeurs de x observées dans l'intervalle $(x_i - \frac{h}{2}, x_i + \frac{h}{2})$,
quelle que soit y .

De la même façon,

$$A_y = \iint_D y T_{xy} d.c dy, \quad (29)$$

$$\sigma_y^2 = \iint_D (y - A_y)^2 T_{xy} dx dy. \quad (30)$$

et, en pratique,

$$\begin{aligned} A_y &= \frac{1}{N} \sum_i \sum_j y_j \psi(x_i, y_j) h \cdot k \\ &= \frac{1}{N} \sum_j y_j \pi(y_j) k \end{aligned} \quad (31)$$

et

$$\begin{aligned} \sigma_y^2 &= \frac{1}{N} \sum_i \sum_j (y_i - A_y)^2 \psi(x_i, y_j) h \cdot k \\ &= \frac{1}{N} \sum_j (y_j - A_y)^2 \pi(y_j) k, \end{aligned} \quad (32)$$

en posant $\pi(y_j) = \sum_i \psi(x_i, y_j) h$.

Outre les écarts quadratiques σ_x^2 et σ_y^2 , s'introduit, par la présence des deux variables x et y , un « moment » d'une autre espèce, qu'on a appelé l'écart rectangulaire :

$$E_{xy}^2 = \iint_D (x - A_x)(y - A_y) T_{xy} dx dy, \quad (33)$$

que nous calculerons par la relation

$$E_{xy}^2 = \frac{1}{N} \sum_i \sum_j (x_i - A_x)(y_j - A_y) \psi(x_i, y_j) h \cdot k. \quad (34)$$

L'exposant rappelle que E_{xy}^2 est une grandeur de même nature ou de même « dimension », au sens de la physique, que σ_x^2 ou σ_y^2 . E_{xy}^2 est donc la valeur moyenne du produit $(x - A_x)(y - A_y)$. Il peut être positif, nul ou négatif et, en valeur absolue,

$$|E_{xy}^2| \leq \sigma_x \cdot \sigma_y,$$

ou bien

$$\frac{E_{xy}^2}{\sigma_x \cdot \sigma_y} \leq 1.$$

Autrement dit, le rapport $\rho_{yx} = \frac{E_{xy}^2}{\sigma_x \cdot \sigma_y}$ [35] peut prendre, à priori, toute valeur comprise entre -1 et $+1$. Ce rapport est le coefficient de corrélation des variables statistiques x et y .

Supposons que σ_x et σ_y ne soient pas nuls. Dans ce cas, $\rho_{xy} = 0$ revient à $E_{xy}^2 = 0$. On peut démontrer alors que, si les variables x et y sont indépendantes,

E^2_{xy} et le coefficient de corrélation de xy sont nuls; mais la réciproque n'est pas toujours vraie. Elle l'est dans un nombre élevé de cas particuliers, à vrai dire les plus fréquents.

De plus, on démontre que si $|\rho_{xy}|=1$, il existe une relation fonctionnelle, linéaire, entre x et y du type $ax + by + c = 0$, et réciproquement.

Plus exactement, cette relation s'écrit, dans ce cas,

$$y - A_y = \frac{\sigma_y}{\sigma_x} (x - A_x). \quad (36)$$

Pour toutes les valeurs de ρ_{xy} , telles que

$$0 < |\rho_{xy}| \leq 1,$$

les variables x et y sont dépendantes au sens des probabilités.

§ 15. Signification du coefficient de corrélation.

Les conclusions qu'on peut tirer du calcul de ρ_{xy} , au moyen de la formule de définition [35] et des formules [28], [32] et [34], ne sont précises que dans un cas : celui où $|\rho_{xy}|=1$ ou a une valeur si voisine de 1 qu'on peut estimer, en raison des écarts possibles, qu'il en est ainsi. Dans ce cas, x et y sont liés par une relation linéaire.

Si $\rho_{xy}=0$, ou si $0 < |\rho_{xy}| < 1$, on ne peut tirer aucune conclusion précise.

Parler de corrélation positive lorsque $0 < \rho_{xy} \leq 1$, et de corrélation négative lorsque $-1 \leq \rho_{xy} < 0$, est affaire de définition; parler de corrélation forte, lorsque ρ_{xy} est plus près de 1 que de 0, et de corrélation faible dans le cas inverse, est d'une imprécision totale.

Encore une fois, le problème général qui se pose ici est celui de l'ajustement, plus complexe, de lois de répartition à deux variables statistiques.

§ 16. Une loi particulière d'ajustement de répartition à deux variables : la loi de BRAVAIS.

On a ainsi exposé d'une façon sommaire et générale les bases de l'étude des distributions de couples de variables statistiques. Il en ressort que toute loi de probabilité à deux variables statistiques, bien définie par la donnée du taux T_{xy} et du domaine certain d'observation D , donne lieu à des valeurs précises des écarts quadratiques et rectangulaires, donc à une valeur du coefficient de corrélation ρ_{yx} .

Inversement, le coefficient de corrélation ne caractérise pas une distribution de couples (x, y) , à moins que, pratiquement, on ne trouve que ρ_{xy} est nul ou égal à un, en valeur absolue. Si l'on a une tendance à donner au coefficient de corrélation une signification plus large que celle qui lui est accordée ici, c'est que, dans certains cas particuliers, la loi de distribution peut être représentée au

moyen d'une fonction à deux variables, où le coefficient ρ_{xy} apparaît explicitement.

La loi, généralement représentée par des fonctions de ce type, est la loi de BRAVAIS :

$$T_{xy} = \frac{1}{2\pi\sigma_x\sigma_y} \cdot \frac{1}{\sqrt{1-\rho_{xy}^2}} e^{-\frac{1}{2(1-\rho_{xy}^2)} \left[\frac{(x-A_x)^2}{\sigma_x^2} - \frac{2\rho_{xy}(x-A_x)(y-A_y)}{\sigma_x\sigma_y} + \frac{(y-A_y)^2}{\sigma_y^2} \right]}. \quad (37)$$

Entre la loi de BRAVAIS et les lois plus générales fournies par des méthodes d'ajustement existent des liens analogues à ceux existant entre la loi de GAUSS et les lois générales des distributions de la fréquence d'une variable statistique, obtenues par les méthodes de PEARSON ou des séries d'HERMITE.

Cette loi s'applique avec succès dans les cas où les mesures portent sur des couples de variables x, y dont chacune suit une loi de répartition rappelant la loi de GAUSS. Il en serait ainsi pour un des exemples dont nous avons parlé : celui de la mesure des longueur et largeur d'œufs de poule d'une espèce. Mais dans l'étude qui nous occupe, la loi de BRAVAIS est inapplicable, car les caractères x et y sont répartis, séparément, de façon tout à fait quelconque, puisque ce sont des mesures de longueurs sur des individus pris au hasard, à tous les stades de leur développement.

Il est aisé de démontrer que les lignes de régression de cette loi sont des droites. La droite de régression de x en y a pour équation

$$y - A_y = \rho_{xy} \frac{\sigma_y}{\sigma_x} (x - A_x), \quad (38)$$

et la droite de régression de x en y ,

$$x - A_x = \rho_{xy} \frac{\sigma_x}{\sigma_y} (y - A_y). \quad (39)$$

§ 17. Droites de régression d'une distribution quelconque.

Les lignes de régression de y en x et de x en y :

$$\begin{aligned} y &= A_{y:x}, \\ x &= A_{x:y}, \end{aligned}$$

sont donc, en général, distinctes des droites d'équations [38] et [39].

Néanmoins, ces droites, quelle que soit la loi de distribution des couples (x, y) , conservent une signification. La droite de régression de y en x est une droite telle que la somme des carrés des différences entre les ordonnées observées de y et celles correspondantes sur la droite soit minimum. La propriété est analogue pour la droite de régression de x en y .

Les équations de ces droites peuvent être obtenues en utilisant une méthode générale, dite des moindres carrés, sur laquelle on ne peut s'étendre ici [cfr.

R. DELTHEIL ⁽⁵⁹⁾]. Cette méthode permet de trouver les équations [YULE, G. U., et KENDALL, M. ⁽⁶⁰⁾], de forme analytique quelconque, de toutes les courbes ayant cette propriété. Parmi ces courbes figurent notamment les lignes de régression elles-mêmes.

De plus, si, pratiquement $\rho_{xy}=1$, les deux équations [38] et [39] sont identiques et expriment la dépendance linéaire fonctionnelle qui existe entre x et y :

$$y - A_y = \frac{\sigma_y}{\sigma_x} (x - A_x). \quad (40)$$

§ 18. Erreur systématique et écart quadratique du coefficient de corrélation.

En toute rigueur, on a

$$\rho_{xy} = \frac{E_{xy}^2}{\sigma_x \cdot \sigma_y},$$

ou bien

$$\rho_{xy} = \frac{\iint_D (x - A_x)(y - A_y) T_{xy} dx dy}{\iint_D (x - A_x)^2 T_{xy} dx dy \cdot \iint_D (y - A_y)^2 T_{xy} dx dy},$$

T_{xy} étant le taux de probabilité de x et y .

En pratique, le coefficient de corrélation est calculé empiriquement, en utilisant les résultats bruts d'une expérience portant sur N couples de valeurs des variables (x, y) . Aux écarts rectangulaires et quadratiques théoriques, on substitue des écarts qui seront appelés empiriques. On peut établir des formules de correction des écarts, donnant la valeur des écarts théoriques en fonction des écarts empiriques. Les différences entre écarts théoriques et empiriques constituent des erreurs systématiques sur ces écarts, au sens habituel qu'on accorde à cette expression, dans la théorie des erreurs. Il en résulte une erreur systématique sur le coefficient de corrélation qui s'exprime sous une forme générale rendant les calculs extrêmement longs. Cette forme se simplifie considérablement si la distribution des couples (x, y) suit une loi « normale » de BRAVAIS. On peut alors considérer que l'erreur systématique du coefficient de corrélation est l'erreur par défaut

$$E = \frac{1 - \rho_{xy}^2}{2N} \cdot \rho_{xy},$$

qu'il convient d'ajouter au coefficient empirique de corrélation pour obtenir de façon très approchée le coefficient théorique [R. RISSER et C. E. TRAYNARD ⁽⁶¹⁾].

⁽⁵⁹⁾ DELTHEIL, R., 1930, pp. 99-103, 115-134.

⁽⁶⁰⁾ YULE, G. U. et KENDALL, M., 1945, pp. 209-210, 311-313.

⁽⁶¹⁾ RISSER, R. et TRAYNARD, C. E., 1933, pp. 195-197.

De même, l'écart quadratique du coefficient de corrélation, qui est la valeur moyenne du carré des écarts entre les coefficients de corrélation empirique (considéré comme une variable statistique) et la valeur moyenne du coefficient de corrélation empirique s'exprime également sous une forme générale assez rebutante.

En première approximation, dans le cas d'une distribution suivant une loi de BRAVAIS, on a

$$\sigma_{\rho_{xy}}^2 = \frac{(1 - \rho_{xy}^2)^2}{N},$$

ou bien

$$\sigma_{\rho_{xy}} = \frac{1 - \rho_{xy}^2}{\sqrt{N}}.$$

Si la distribution des couples (x, y) n'obéit pas à la loi de BRAVAIS, cette dernière expression fixe avec une approximation très satisfaisante l'ordre de grandeur de l'écart quadratique.

Ce qu'il faut en retenir, d'une manière générale, c'est que l'erreur systématique est proportionnelle à $\frac{1}{N}$ et l'écart quadratique à $\frac{1}{\sqrt{N}}$. Pour autant que N soit suffisamment grand, l'erreur systématique devient négligeable devant l'écart quadratique.

Par exemple, supposons $\rho_{xy} = 0,9$. On trouve successivement :

a) Pour $N = 150$:

$$\sigma_{\rho} = \frac{0,19}{12,24} = 0,0155,$$

$$E = \frac{0,19}{300} \times 0,9 = 0,0006;$$

b) Pour $N = 225$:

$$\sigma_{\rho} = \frac{0,19}{15} = 0,0127,$$

$$E = \frac{0,19}{450} \times 0,9 = 0,0004;$$

c) Pour $N = 400$:

$$\sigma_{\rho} = \frac{0,19}{20} = 0,0095,$$

$$E = \frac{0,19}{800} \times 0,9 = 0,0002.$$

On voit que l'erreur systématique affecte tout au plus la troisième décimale du coefficient de corrélation, puisqu'elle est inférieure au 0,001. Encore s'agit-il ici d'expériences portant sur des nombres assez réduits d'individus. Comme l'erreur systématique du coefficient de corrélation est une erreur par défaut, on

peut en tenir compte en conservant trois décimales du coefficient de corrélation empirique, la dernière étant forcée.

Par contre, l'écart quadratique du coefficient de corrélation affecte très souvent la deuxième décimale et l'on peut admettre que le coefficient théorique soit situé dans un intervalle dont l'ordre de grandeur est 0,03, centré sur la valeur empirique obtenue au moyen de nos procédés de calcul. Cet ordre de grandeur n'a rien d'absolu; il convient au problème particulier que nous avons à traiter.

C. — MENSURATIONS DES INDIVIDUS.

Parmi les gisements ayant livré des *Anthraconauta belgica* cités dans la liste aux pages 11 et 12, deux d'entre eux ont fourni un matériel suffisamment abondant et en assez bon état de conservation pour se prêter à des mensurations (voir planches A et B. Ce sont : Pl. Chenée, Angleur (5), Pl. Andenne, Bois de Stud; ils seront désignés par la suite : A. — Angleur et A. — Stud. Pour la comparaison des espèces provenant de ces gisements et attribués à *A. belgica*, avec des espèces d'*A. minima*, il a été examiné des matériaux de cette dernière provenant du charbonnage de Wérister, siège José, étage de 242 m., Bacnure Nord, stampe des veines Venta à Beaujardin, première veinette sous Grosse, niveau n° 43. Ils seront désignés par la suite : A. — Wérister.

Les mesures effectuées suivant les méthodes indiquées pages 13 et 14 ont porté sur un nombre d'individus variant de 421 à 433 pour A. — Wérister, 193 à 209 pour A. — Angleur, 146 à 150 pour A. — Stud, suivant les caractères mesurés.

CHAPITRE III.

D. — ÉTUDE DE LA CORRELATION DES CARACTÈRES.

INTRODUCTION. — Il a été vu au chapitre II que la notion de corrélation permet d'étudier la variabilité des caractères mesurés, en fonction du développement (ou âge) de ces individus. Il a été exposé la signification de la corrélation et du coefficient par lequel on peut en donner une mesure, imparfaite d'ailleurs, sauf dans les cas extrêmes où il est nul ou égal à ± 1 . Dans ce dernier cas, on peut conclure, en toute rigueur, à une liaison fonctionnelle linéaire entre les caractères rapprochés.

Il est évident — et l'attention a déjà été attirée sur ce sujet — qu'une distribution expérimentale de couples de variables statistiques, dont on sait à l'avance qu'elles sont liées par une relation linéaire, ne donnera un coefficient égal à ± 1 que de façon tout à fait exceptionnelle. En statistique, l'écart est la règle et non l'exception. Dès lors, compte tenu des écarts, obtenir un coefficient

peut en tenir compte en conservant trois décimales du coefficient de corrélation empirique, la dernière étant forcée.

Par contre, l'écart quadratique du coefficient de corrélation affecte très souvent la deuxième décimale et l'on peut admettre que le coefficient théorique soit situé dans un intervalle dont l'ordre de grandeur est 0,03, centré sur la valeur empirique obtenue au moyen de nos procédés de calcul. Cet ordre de grandeur n'a rien d'absolu; il convient au problème particulier que nous avons à traiter.

C. — MENSURATIONS DES INDIVIDUS.

Parmi les gisements ayant livré des *Anthraconauta belgica* cités dans la liste aux pages 11 et 12, deux d'entre eux ont fourni un matériel suffisamment abondant et en assez bon état de conservation pour se prêter à des mensurations (voir planches A et B. Ce sont : Pl. Chenée, Angleur (5), Pl. Andenne, Bois de Stud; ils seront désignés par la suite : A. — Angleur et A. — Stud. Pour la comparaison des espèces provenant de ces gisements et attribués à *A. belgica*, avec des espèces d'*A. minima*, il a été examiné des matériaux de cette dernière provenant du charbonnage de Wérister, siège José, étage de 242 m., Bacnure Nord, stampe des veines Venta à Beaujardin, première veinette sous Grosse, niveau n° 43. Ils seront désignés par la suite : A. — Wérister.

Les mesures effectuées suivant les méthodes indiquées pages 13 et 14 ont porté sur un nombre d'individus variant de 421 à 433 pour A. — Wérister, 193 à 209 pour A. — Angleur, 146 à 150 pour A. — Stud, suivant les caractères mesurés.

CHAPITRE III.

D. — ÉTUDE DE LA CORRELATION DES CARACTÈRES.

INTRODUCTION. — Il a été vu au chapitre II que la notion de corrélation permet d'étudier la variabilité des caractères mesurés, en fonction du développement (ou âge) de ces individus. Il a été exposé la signification de la corrélation et du coefficient par lequel on peut en donner une mesure, imparfaite d'ailleurs, sauf dans les cas extrêmes où il est nul ou égal à ± 1 . Dans ce dernier cas, on peut conclure, en toute rigueur, à une liaison fonctionnelle linéaire entre les caractères rapprochés.

Il est évident — et l'attention a déjà été attirée sur ce sujet — qu'une distribution expérimentale de couples de variables statistiques, dont on sait à l'avance qu'elles sont liées par une relation linéaire, ne donnera un coefficient égal à ± 1 que de façon tout à fait exceptionnelle. En statistique, l'écart est la règle et non l'exception. Dès lors, compte tenu des écarts, obtenir un coefficient

dont la valeur absolue est de l'ordre de 0,90 est un résultat qui permet de conclure à une dépendance quasi linéaire, surtout si l'on dispose de collections relativement peu nombreuses.

MÉTODES. — Les calculs qui suivent ont été effectués selon les principes et méthodes énoncés précédemment.

Afin de calculer les coefficients de corrélation, des tables de corrélation à double entrée ont été établies. Pour arriver à ce but, il a fallu :

1° Partager en intervalles partiels de longueurs convenables les intervalles de variation des deux variables mises chaque fois en cause.

2° Compter les fréquences d'apparition des couples (x, y) dans chacun des rectangles ainsi obtenus, fréquences désignées au chapitre II par la notation $\varphi(x_i, y_j)hk$. On obtient ainsi un relevé statistique brut de la distribution (quelconque) de fréquence des couples (x, y) .

3° Calculer les moyennes liées de l'un des caractères par rapport à l'autre.

Pour préparer le calcul définitif des coefficients de corrélation et de régression, il a été calculé successivement dans chaque cas :

- a) les valeurs moyennes des caractères mesurés;
- b) les écarts quadratiques ou la dispersion des caractères;
- c) les écarts rectangulaires, valeurs moyennes des produits des écarts à la moyenne.

Voici le détail de toute la suite des opérations conduisant à l'établissement de ces valeurs :

Pour chaque population, la corrélation entre les caractères d et a , d et b , d et g a été étudiée, le caractère d étant choisi comme élément de référence en tant que mesure du diamètre de croissance maximum de la coquille.

On a désigné par d_0 ou $A(d)$, a_0 ou $A(a)$, b_0 ou $A(b)$, g_0 ou $A(g)$ les moyennes des caractères mesurés; par σ_d , σ_a , σ_b , σ_g , leurs écarts quadratiques; par E_{da}^2 , E_{db}^2 , E_{dg}^2 les écarts rectangulaires. Les coefficients de corrélation se définissent alors par

$$\rho_{da} = \frac{E_{da}^2}{\sigma_d \cdot \sigma_a},$$

$$\rho_{db} = \frac{E_{db}^2}{\sigma_d \cdot \sigma_b},$$

$$\rho_{dg} = \frac{E_{dg}^2}{\sigma_d \cdot \sigma_g}.$$

Il a paru nécessaire de donner quelques indications au sujet des nombres figurant dans les nombreux tableaux qui suivent.

Dans les tables de corrélation, chacune des colonnes, portant en tête l'indication d'un intervalle partiel de variation de d et, entre parenthèses, l'écart $d-d_0$ du centre de l'intervalle à la moyenne, fournit la répartition « liée » de l'autre variable a , b ou g . Au bas de la colonne est noté le total des observations de valeurs de d de l'intervalle partiel susdit et la moyenne liée de a , b ou g en d .

Chacune des lignes donne les renseignements analogues, sauf à renverser le rôle des variables.

Ces tableaux de calcul des écarts quadratiques ou dispersions se passent de commentaires, de même que les tableaux de calcul des écarts rectangulaires reproduisant, sous une forme différente, les tables de corrélation.

REPRÉSENTATION GRAPHIQUE. — Les points expérimentaux, dont les coordonnées sont définies par les moyennes liées et qui, théoriquement et de façon générale, sont situés sur les lignes de régression, ont été représentés géométriquement sur des graphiques. On y a tracé les droites de régression et la droite idéale de la relation fonctionnelle linéaire qui correspondrait, dans chaque cas, à $\rho=1$ (Pl. I à IX).

§ 1. TABLEAUX DE CORRÉLATION.

Gisement de Wérister

[*Anthraconauta minima* HIND (non LUDWIG), espèce de référence.]

1. Tableau de corrélation entre les caractères d et a .

a ($a - a_0$)	d ($d - d_0$)														Fréquences a	Moyennes liées d en a
	1-2 (-4,81)	2-3 (-3,81)	3-4 (-2,81)	4-5 (-1,81)	5-6 (-0,81)	6-7 (+0,19)	7-8 (+1,19)	8-9 (+2,19)	9-10 (+3,19)	10-11 (+4,19)	11-12 (+5,19)	12-13 (+6,19)	13-14 (+7,19)	14-15 (+8,19)		
1-2 (-3,11)	3	9	1	13	2,35
2-3 (-2,11)	.	2	40	15	1	58	3,76
3-4 (-1,11)	.	.	9	60	40	5	1	115	4,88
4-5 (-0,11)	.	.	.	6	30	33	11	80	6,11
5-6 (+0,89)	2	25	34	18	2	81	7,41
6-7 (+1,89)	1	10	19	12	7	1	.	.	.	50	8,84
7-8 (+2,89)	10	10	1	.	.	.	21	10,07
8-9 (+3,89)	1	4	9	.	.	.	14	11,07
9-10 (+4,89)	1	1	.	.	2	12
10-11 (+5,89)
11-12 (+6,89)	1	1	14,5
Fréquences d	3	11	50	81	73	64	56	37	25	21	12	1	.	1	435	.
Moyennes liées a en d	1,50	1,70	2,66	3,39	3,95	4,84	5,45	6,01	6,90	7,36	8,33	9,50	.	11,5	.	.

2. Tableau de corrélation entre les caractères d et b .

b ($b - b_0$)	d ($d - d_0$)														Fréquences b	Moyennes liées d en b
	1-2 (-4,81)	2-3 (-3,81)	3-4 (-2,81)	4-5 (-1,81)	5-6 (-0,81)	6-7 (+0,19)	7-8 (+1,19)	8-9 (+2,19)	9-10 (+3,19)	10-11 (+4,19)	11-12 (+5,19)	12-13 (+6,19)	13-14 (+7,19)	14-15 (+8,19)		
1-2 (-4,94)	2	2	2
2-3 (-3,94)	1	7	1	1	10	2,70
3-4 (-2,94)	.	4	33	3	40	3,48
4-5 (-1,94)	.	.	16	56	12	1	85	4,47
5-6 (-0,94)	.	.	.	21	45	11	2	1	80	5,46
6-7 (+0,06)	16	32	6	1	55	6,35
7-8 (1,06)	18	32	10	60	7,37
8-9 (2,06)	2	16	18	7	43	8,19
9-10 (3,06)	6	12	7	.	.	.	25	9,54
10-11 (4,06)	1	6	9	1	.	.	17	10,09
11-12 (5,06)	3	1	.	.	4	10,75
12-13 (6,06)	2	10	1	.	13	11,42
13-14 (7,06)
14-15 (8,06)	1	1	14,5
Fréquences d	3	11	50	81	73	64	56	37	25	21	12	1	.	1	435	.
Moyennes liées b en d	1,83	2,86	3,80	4,70	5,55	6,64	7,61	8,31	9,46	10,50	12,25	12,50	.	14,50	.	.

3. Tableau de corrélation entre les caractères d et g .

g ($g - g_0$)	d ($d - d_0$)														Fréquences g	Moyennes liées d en g
	1-2 (-4,81)	2-3 (-3,81)	3-4 (-2,81)	4-5 (-1,81)	5-6 (-0,81)	6-7 (+0,19)	7-8 (+1,19)	8-9 (+2,19)	9-10 (+3,19)	10-11 (+4,19)	11-12 (+5,19)	12-13 (+6,19)	13 14 (+7,19)	14-15 (+8,19)		
1-2 (-2,13)	3	11	24	1	39	3,09
2-3 (-1,13)	.	.	25	64	27	1	117	4,53
3-4 (-0,13)	.	.	1	10	42	45	19	2	119	6,15
4-5 (+0,87)	.	.	.	6	4	18	33	26	4	2	93	7,46
5-6 (+1,87)	4	9	17	13	4	.	.	.	47	9,58
6-7 (+2,87)	4	6	5	.	.	.	15	10,57
7-8 (+3,87)	3	1	.	.	4	11,75
8-9 (+4,87)	1	1	14,50
Fréquences d	3	11	50	81	73	64	56	37	25	21	12	1	.	1	435	.
Moyennes liées g en d	1,50	1,50	2,04	2,76	3,18	3,77	4,23	4,69	5,50	5,69	6,42	7,50	.	8,50	.	.

Gisement d'Angleur
[*Anthraconauta belgica* Auctorum.]

1. Tableau de corrélation entre les caractères d et a .

a ($a - a_0$)	d ($d - d_0$)									Fréquences a	Moyennes liées d en a
	1-3 (-4,11)	3-5 (-2,11)	5-7 (-0,11)	7-9 (+1,89)	9-11 (+3,89)	11-13 (+5,89)	13-15 (+7,89)	15-17 (+9,89)	17-19 (+11,89)		
1-3 (-2,63)	14	26	40	3,30
3-5 (-0,63)	.	46	35	4	1	86	5,07
5-7 (1,37)	.	.	15	41	10	1	.	.	.	67	7,91
7-9 (3,37)	.	.	.	4	6	.	1	.	.	11	9,64
9-11 (5,37)	1	1	.	1	1	4	14,00
11-13 (7,37)	0	—
13-15 (9,37)	1	.	1	16,00
Fréquences a	14	72	50	49	18	2	1	2	1	209	—
Moyennes liées a en d	2	3,42	4,60	6	6,78	8	8	12	10	.	—

2. Tableau de corrélation entre les caractères d et b .

b ($b - b_0$)	d ($d - d_0$)									Fréquences b	Moyennes liées d en b
	1-3 (-4,11)	3-5 (-2,11)	5-7 (-0,11)	7-9 (+1,89)	9-11 (+3,89)	11-13 (+5,89)	13-15 (+7,89)	15-17 (+9,89)	17-19 (+11,89)		
1-3 (-4,45)	7	3	10	2,60
3-5 (-2,45)	7	59	2	68	3,85
5-7 (-0,45)	.	10	38	5	53	5,81
7-9 (1,55)	.	.	10	28	2	40	7,60
9-11 (3,55)	.	.	.	16	9	26	8,72
11-13 (5,55)	7	2	1	.	.	10	10,80
13-15 (7,55)	1	.	1	16,00
15-17 (9,55)	1	1	18,00
17-19 (11,55)	1	.	1	16,00
Fréquences d	14	72	50	49	18	2	1	2	1	209	.
Moyennes liées b en d	3	4,19	6,32	8,45	10,56	12	12	16	16	.	.

3. Tableau de corrélation entre les caractères d et g .

g ($g - g_0$)	d ($d - d_0$)									Fréquences g	Moyennes liées d en g
	1-3 (-4,08)	3-5 (-2,08)	5-7 (-0,08)	7-9 (+1,92)	9-11 (+3,92)	11-13 (+5,92)	13-15 (+7,92)	15-17 (+9,92)	17-19 (+11,92)		
1-2 (-2,39)	11	10	21	2,95
2-3 (-1,39)	4	41	5	1	51	4,12
3-4 (-0,39)	.	14	25	4	1	44	5,54
4-5 (+0,61)	.	3	14	17	4	38	7,16
5-6 (1,61)	.	2	2	20	4	28	7,86
6-7 (2,61)	.	.	.	5	6	11	9,09
7-8 (3,61)	.	.	.	1	2	1	1	.	.	5	10,80
8-9 (4,61)	1	.	.	.	1	12,00
9-10 (5,61)	1	.	1	16,00
10-11 (6,61)	1	1	2	17,00
Fréquences d	15	70	46	48	17	2	1	2	1	202	.
Moyennes liées g en d	1,77	2,73	3,78	5,06	5,73	8	7,50	10	10,50	.	.

Gisement de Stud
[*Anthraconauta belgica* Auctorum.]

1. Tableau de corrélation entre les caractères d et a .

a ($a - a_0$)	d ($d - d_0$)											Fréquences a	Moyennes liées d en a
	2,5-3,5 (-4,37)	3,5-4,5 (-3,37)	4,5-5,5 (-2,37)	5,5-6,5 (-1,37)	6,5-7,5 (-0,37)	7,5-8,5 (+0,63)	8,5-9,5 (+1,63)	9,5-10,5 (+2,63)	10,5-11,5 (+3,63)	11,5-12,5 (+4,63)	12,5-13,5 (+5,63)		
2,5-3,5 (-2,36)	6	11	5	2	24	4,13
3,5-4,5 (-1,26)	.	2	12	9	4	1	28	5,64
4,5-5,5 (-0,36)	.	.	2	3	10	14	1	30	7,30
5,5-6,5 (+0,64)	.	.	.	1	8	8	4	2	.	.	.	23	7,91
6,5-7,5 (1,64)	2	6	9	5	7	.	.	29	9,31
7,5-8,5 (2,64)	4	3	2	2	11	11,18
8,5-9,5 (3,64)	2	1	1	4	11,75
Fréquences d	6	13	19	15	24	29	14	11	12	3	3	149	.
Moyennes liées a en d	3	3,15	3,84	4,20	5,33	5,66	6,57	7,18	7,58	8,33	8,33	.	.

2. Tableau de corrélation entre les caractères d et b .

b ($b - b_0$)	d ($d - d_0$)											Fréquences b	Moyennes liées d en b
	2,5-3,5 (-4,37)	3,5-4,5 (-3,37)	4,5-5,5 (-2,37)	5,5-6,5 (-1,37)	6,5-7,5 (-0,37)	7,5-8,5 (+0,63)	8,5-9,5 (+1,63)	9,5-10,5 (+2,63)	10,5-11,5 (+3,63)	11,5-12,5 (+4,63)	12,5-13,5 (+5,63)		
2,5-3,5 (-4,56)	2	1	3	3,33
3,5-4,5 (-3,56)	3	4	2	1	10	4,10
4,5-5,5 (-2,56)	1	8	10	3	22	4,68
5,5-6,5 (-1,56)	.	.	6	9	4	1	20	6,00
6,5-7,5 (-0,56)	.	.	1	1	12	3	17	7,00
7,5-8,5 (+0,44)	.	.	.	1	7	17	5	30	7,87
8,5-9,5 (1,44)	1	7	7	3	1	.	.	19	8,79
9,5-10,5 (2,44)	1	2	2	2	.	.	7	9,71
10,5-11,5 (3,44)	5	4	1	.	10	10,60
11,5-12,5 (4,44)	1	3	2	1	7	11,43
12,5-13,5 (5,44)	2	.	1	3	11,67
13,5-14,5 (6,44)
14,5-15,5 (7,44)	1	1	13,00
Fréquences d	6	13	19	15	24	29	14	11	12	3	3	149	.
Moyennes liées b en d	3,83	4,54	5,32	5,87	7,21	8,14	8,79	10,36	11,25	11,67	13,33	.	.

3. Tableau de corrélation entre les caractères d et g .

g ($g - g_0$)	d ($d - d_0$)											Fréquences g	Moyennes liées d en g
	2,5-3,5 (-4,37)	3,5-4,5 (-3,37)	4,5-5,5 (-2,37)	5,5-6,5 (-1,37)	6,5-7,5 (-0,37)	7,5-8,5 (+0,63)	8,5-9,5 (+1,63)	9,5-10,5 (+2,63)	10,5-11,5 (+3,63)	11,5-12,5 (+4,63)	12,5-13,5 (+5,63)		
1,5-2,5 (-2,44)	5	5	2	1	13	3,92
2,5-3,5 (-1,44)	1	7	13	7	1	1	30	5,10
3,5-4,5 (-0,44)	.	1	4	6	13	6	30	6,63
4,5-5,5 (+0,56)	.	.	.	1	8	20	11	3	2	.	.	45	8,29
5,5-6,5 (1,56)	2	2	3	5	5	2	.	19	9,79
6,5-7,5 (2,56)	3	4	.	2	9	11,11
7,5-8,5 (3,56)	1	1	1	3	12,00
Fréquences d	6	13	19	15	24	29	14	11	12	3	3	149	.
Moyennes liées g en d	2,17	2,69	3,10	3,47	4,46	4,79	5,21	6	6,33	6,67	7,33	.	.

§ 2. CALCUL DES MOYENNES.

Gisement de Wérister.

[*Anthraconauta minima* HIND (non LUDWIG), espèce de référence.]

$$d_0 = \Lambda(d) = \frac{1}{435} [217,5 + 3 + 22 + 150 + 324 + 365 + 384 + 392 + 296 \\ + 225 + 210 + 132 + 12 + 14] = 6,31$$

$$a_0 = \Lambda(a) = \frac{1}{435} [217,5 + 13 + 116 + 345 + 320 + 405 + 300 + 147 + 112 + 18 + 11] = 4,61$$

$$b_0 = \Lambda(b) = \frac{1}{435} [217,5 + 2 + 20 + 120 + 340 + 400 + 330 + 420 + 344 \\ + 225 + 170 + 44 + 156 + 14] = 6,44$$

$$g_0 = \Lambda(g) = \frac{1}{435} [217,5 + 39 + 234 + 357 + 372 + 235 + 90 + 28 + 8] = 3,63$$

Gisement d'Angleur.

[*Anthraconauta belgica* Auctorum.]

$$d_0 = \Lambda(d) = \frac{1}{209} [28 + 288 + 300 + 392 + 180 + 24 + 14 + 32 + 18] = 6,11 \quad (1)$$

$$d_0 = \Lambda(d) = \frac{1}{202} [30 + 280 + 276 + 384 + 170 + 24 + 14 + 32 + 18] = 6,08 \quad (2)$$

$$a_0 = \Lambda(a) = \frac{1}{209} [80 + 344 + 402 + 88 + 40 + 14] = 4,63$$

$$b_0 = \Lambda(b) = \frac{1}{209} [20 + 272 + 318 + 320 + 250 + 120 + 14 + 16 + 18] = 6,45$$

$$g_0 = \Lambda(g) = \frac{1}{202} [101 + 21 + 102 + 132 + 152 + 140 + 66 + 35 + 8 + 9 + 20] = 3,89 \quad (3)$$

Gisement de Stud.

[*Anthraconauta belgica* Auctorum.]

$$d_0 = \Lambda(d) = \frac{1}{149} [18 + 52 + 95 + 90 + 168 + 232 + 126 + 110 + 132 + 36 + 39] = 7,37$$

$$a_0 = \Lambda(a) = \frac{1}{149} [72 + 112 + 150 + 138 + 203 + 88 + 36] = 5,36$$

$$b_0 = \Lambda(b) = \frac{1}{149} [9 + 40 + 110 + 120 + 119 + 240 + 171 + 70 + 110 + 84 + 39 + 15] = 7,56$$

$$g_0 = \Lambda(g) = \frac{1}{149} [26 + 90 + 120 + 225 + 114 + 63 + 24] = 4,44$$

(1) Sur 209 individus.

(2) et (3) Sur 202 individus.

§ 3. CALCUL DES DISPERSIONS.

Gisement de Wérister.

[*Anthraconauta minima* HIND (non LUDWIG), espèce de référence.]

Calcul de σ_d

d	d^2	Fréquences	$F \times d^2$
1,5	2,25	3	6,75
2,5	6,25	11	68,75
3,5	12,25	50	612,50
4,5	20,25	81	1.640,25
5,5	30,25	73	2.208,25
6,5	42,25	64	2.704,—
7,5	56,25	56	3.150,—
8,5	72,25	37	2.673,25
9,5	90,25	25	2.256,25
10,5	110,25	21	2.315,25
11,5	132,25	12	1.587,—
12,5	156,25	1	156,25
13,5	182,25	—	—
14,5	210,25	1	210,25
		435	$\Sigma = 19.588,75$

$$A(d^2) = 45,0316$$

$$[A(d)]^2 = 39,8161$$

$$\sigma_d^2 = 5,2155$$

$$\sigma_d = 2,28$$

Calcul de σ_a

a	a^2	Fréquences	$F \times a^2$
1,5	2,25	13	29,25
2,5	6,25	58	362,50
3,5	12,25	115	1.408,75
4,5	20,25	80	1.620,—
5,5	30,25	81	2.450,25
6,5	42,25	50	2.112,50
7,5	56,25	21	1.181,25
8,5	72,25	14	1.011,50
9,5	90,25	2	180,50
10,5	110,25	—	—
11,5	132,25	1	132,25
		435	$\Sigma = 10.488,75$

$$A(a^2) = 24,1121$$

$$[A(a)]^2 = 21,2521$$

$$\sigma_a^2 = 2,8600$$

$$\sigma_a = 1,69$$

Calcul de σ_b

b	b^2	Fréquences	$F \times b^2$
1,5	2,25	2	4,5
2,5	6,25	10	62,5
3,5	12,25	40	490,—
4,5	20,25	85	1.721,25
5,5	30,25	80	2.420,—
6,5	42,25	55	2.323,75
7,5	56,25	60	3.375,—
8,5	72,25	43	3.106,75
9,5	90,25	25	2.256,25
10,5	110,25	17	1.874,25
11,5	132,25	4	529,—
12,5	156,25	13	2.031,25
13,5	182,25	—	—
14,5	210,25	1	210,25
		435	$\Sigma = 20.404,75$

$$A(b^2) = 46,9126$$

$$[A(b)]^2 = 41,4736$$

$$\sigma_b^2 = 5,4390$$

$$\sigma_b = 2,33$$

Calcul de σ_g

g	g^2	Fréquences	$F \times g^2$
1,5	2,25	33	87,75
2,5	6,25	117	731,25
3,5	12,25	119	1.457,75
4,5	20,25	93	1.883,25
5,5	30,25	47	1.421,75
6,5	42,25	15	633,75
7,5	56,25	4	225,00
8,5	72,25	1	72,25
		435	$\Sigma = 6.512,75$

$$A(g^2) = 14,9701$$

$$[A(g)]^2 = 13,1769$$

$$\sigma_g^2 = 1,7932$$

$$\sigma_g = 1,34$$

Gisement d'Angleur
 [Anthraconauta belgica Auctorum.]

Calcul de $\sigma_d^{(1)}$.

d	d^2	Fréquences	$F \times d^2$
2	4	14	56
4	16	72	1.152
6	36	50	1.800
8	64	49	3.136
10	100	18	1.800
12	144	2	288
14	196	1	196
16	256	2	512
18	324	1	324
		209	$\Sigma = 9.264$

$$A(d^2) = 44,3253$$

$$[A(d)]^2 = 37,3321$$

$$\sigma_d^2 = 6,9932$$

$$\sigma_d = 2,64$$

Calcul de $\sigma_d^{(2)}$.

d	d^2	Fréquences	$F \times d^2$
2	4	15	60
4	16	70	1.120
6	36	46	1.656
8	64	48	3.072
10	100	17	1.700
12	144	2	288
14	196	1	196
16	256	2	512
18	324	1	324
		202	$\Sigma = 8.928$

$$A(d^2) = 44,1980$$

$$[A(d)]^2 = 36,9664$$

$$\sigma_d^2 = 7,2316$$

$$\sigma_d = 2,69$$

(1) Sur 209 individus.

(2) Sur 202 individus.

Calcul de σ_a

a	a^2	Fréquences	$F \times a^2$
2	4	40	160
4	16	86	1.376
6	36	67	2.412
8	64	11	704
10	100	4	400
12	144	—	—
14	196	1	196
		209	$\Sigma = 5.248$

$$A(a^2) = 25,1000$$

$$[A(a)]^2 = 21,4369$$

$$\sigma_a^2 = 3,6631$$

$$\sigma_a = 1,91$$

Calcul de σ_b

b	b^2	Fréquences	$F \times b^2$
2	4	10	40
4	16	68	1.088
6	36	53	1.908
8	64	40	2.560
10	100	25	2.500
12	144	10	1.440
14	196	1	196
16	256	1	256
18	324	1	324
		209	$\Sigma = 10.312$

$$A(b^2) = 49,3397$$

$$[A(b)]^2 = 41,6025$$

$$\sigma_b^2 = 7,7372$$

$$\sigma_b = 2,78$$

Calcul de σ_g

g	g^2	Fréquences	$F \times g^2$
1,5	2,25	21	47,25
2,5	6,25	51	236,25
3,5	12,25	44	539,—
4,5	20,25	38	769,50
5,5	30,25	28	847,—
6,5	42,25	11	464,75
7,5	56,25	5	281,25
8,5	72,25	1	72,25
9,5	90,25	1	90,25
10,5	110,25	2	220,50
		202	$\Sigma = 3.568,00$

$$A(g^2) = 17,6633$$

$$[A(g)]^2 = 15,1321$$

$$\sigma_g^2 = 2,5312$$

$$\sigma_g = 1,59$$

Gisement de Stud.

[*Anthraconauta belgica* Auctorum.]

Calcul de σ_d

d	d^2	Fréquences	$F \times d^2$
3	9	6	54
4	16	13	208
5	25	19	475
6	36	15	540
7	49	24	1.176
8	64	29	1.856
9	81	14	1.134
10	100	11	1.100
11	121	12	1.552
12	144	3	432
13	169	3	507
		149	$\Sigma = 8.934$

$$A(d^2) = 59,9597$$

$$[A(d)]^2 = 54,3169$$

$$\sigma_d^2 = 5,6428$$

$$\sigma_d = 2,37$$

Calcul de σ_a

a	a^2	Fréquences	$F \times a^2$
3	9	24	216
4	16	28	448
5	25	30	750
6	36	23	828
7	49	29	1.421
8	64	11	704
9	81	4	324
		149	$\Sigma = 4.691$

$$A(a^2) = 31,4832$$

$$[A(a)]^2 = 28,7296$$

$$\sigma_a^2 = 2,7536$$

$$\sigma_a = 1,66$$

Calcul de σ_b

b	b^2	Fréquences	$F \times b^2$
3	9	3	27
4	16	10	160
5	25	22	550
6	36	20	720
7	49	17	833
8	64	30	1.920
9	81	19	1.539
10	100	7	700
11	121	10	1.210
12	144	7	1.008
13	169	3	507
14	196	—	—
15	225	1	225
		149	$\Sigma = 9.399$

$$A(b^2) = 63,0805$$

$$[A(b)]^2 = 57,1536$$

$$\sigma_b^2 = 5,9269$$

$$\sigma_b = 2,43$$

Calcul de σ_g

g	g^2	Fréquences	$F \times g^2$
2	4	13	52
3	9	30	270
4	16	30	480
5	25	45	1.125
6	36	19	684
7	49	9	441
8	64	3	192
		149	$\Sigma = 3.244$

$$A(g^2) = 21,7718$$

$$[A(g)]^2 = 19,7136$$

$$\sigma_g^2 = 2,0582$$

$$\sigma_g = 1,43$$

§ 4. — CALCUL DES ÉCARTS RECTANGULAIRES ET DES COEFFICIENTS DE CORRÉLATION ET DE RÉGRESSION.

Gisement de Wérister.

[*Anthraconauta minima* HIND (non LUDWIG), espèce de référence.]

Calcul de l'écart rectangulaire E_{da}^2 et des coefficients ρ_{da} et de régression de a en d et d en a .

$d - d_0$	$a - a_0$	$(d - d_0)(a - a_0)$	Fréquences	$F \times (d - d_0)(a - a_0)$
-4,81	-3,11	14,9591	3	44,8773
-3,81	-3,11	11,8491	9	106,6419
-3,81	-2,11	8,0391	2	16,0782
-2,81	-3,11	8,7391	1	8,7391
-2,81	-2,11	5,9291	40	237,1640
-2,81	-1,11	3,1191	9	28,0719
-1,81	-2,11	3,8191	15	57,2865
-1,81	-1,11	2,0091	60	120,5460
-1,81	-0,11	0,1991	6	1,1946
-0,81	-2,11	1,7091	1	1,7091
-0,81	-1,11	0,8991	40	35,9640
-0,81	-0,11	0,0891	30	2,6730
-0,81	+0,89	-0,7209	2	-1,4418
+0,19	-1,11	-0,2109	5	-1,0545
+0,19	-0,11	-0,0209	33	-0,6897
+0,19	0,89	0,1691	25	4,2275
			A reporter... ..	281

		Report...	281	
+0,19	1,89	0,3591			1	0,3591
1,19	-1,11	-1,3209			1	-1,3209
1,19	-0,11	-0,1309			11	-1,4399
1,19	0,89	1,0591			34	36,0094
1,19	1,89	2,2491			10	22,4910
2,19	0,89	1,9491			18	35,0838
2,19	1,89	4,1391			19	78,6429
3,19	0,89	2,8391			2	5,6782
3,19	1,89	6,0291			12	72,3492
3,19	2,89	9,2191			10	92,1910
3,19	3,89	12,4091			1	12,4091
4,19	1,89	7,9191			7	55,4337
4,19	2,89	12,1091			10	121,0910
4,19	3,89	16,2991			4	65,1964
5,19	1,89	9,8091			1	9,8091
5,19	2,89	14,9991			1	14,9991
5,19	3,89	20,1891			9	181,7019
5,19	4,89	25,3791			1	25,3791
6,19	4,89	30,2691			1	30,2691
8,19	6,89	56,4291			1	56,4291
					435	$\Sigma = 1574,7485$

$$E_{da}^2 = \frac{1}{435} \Sigma = 3,6201$$

$$\sigma_d \cdot \sigma_a = 2,28 \times 1,69 = 3,8532$$

$$\text{Coefficient de corrélation : } \rho_{da} = \frac{E_{da}^2}{\sigma_d \sigma_a} = \frac{3,6201}{3,8532} = 0,940$$

$$\text{Coefficient de régression de } a \text{ en } d : \rho_{da} \frac{\sigma_a}{\sigma_d} = \frac{3,6201}{5,2155} = 0,694$$

$$\text{Coefficient de régression de } d \text{ en } a : \rho_{da} \frac{\sigma_d}{\sigma_a} = \frac{3,6201}{2,8600} = 1,262 = 0,790$$

$$\text{Coefficient angulaire pour } \rho = 1 : \frac{\sigma_a}{\sigma_d} = \frac{1,69}{2,28} = 0,741$$

Calcul de l'écart rectangulaire E^2_{ab} et des coefficients ρ_{ab} et de régression de b en d et d en b .

$\bar{d} - \bar{d}_0$	$\bar{b} - \bar{b}_0$	$(\bar{d} - \bar{d}_0) (\bar{b} - \bar{b}_0)$	Fréquences	$F \times (\bar{d} - \bar{d}_0) (\bar{b} - \bar{b}_0)$
-4,81	-4,94	23,7614	2	47,5228
-4,81	-3,94	18,9514	1	18,9514
-3,81	-3,94	15,0114	7	105,0798
-3,81	-2,94	11,2014	4	44,8056
-2,81	-3,93	11,0714	1	11,0714
-2,81	-2,94	8,2614	33	272,6262
-2,81	-1,94	5,4514	16	87,2224
-1,81	-3,94	7,1314	1	7,1314
-1,81	-2,94	5,3214	3	15,9642
-1,81	-1,94	3,5114	56	196,6384
-1,81	-0,94	1,7014	21	35,7294
-0,81	-1,94	1,5714	12	18,8568
-0,81	-0,94	0,7614	45	34,2630
-0,81	+0,06	-0,0486	16	-0,7776
+0,19	-1,94	-0,3686	1	-0,3686
+0,19	-0,94	-0,1786	11	-1,9646
+0,19	+0,06	0,0114	32	0,3648
+0,19	+1,06	0,2014	18	3,6252
+0,19	+2,06	0,3914	2	0,7828
+1,19	-0,94	-1,1186	2	-2,2372
+1,19	+0,06	0,0714	6	0,4284
1,19	1,06	1,2614	32	40,3648
1,19	2,06	2,4514	16	39,2224
2,19	-0,94	-2,0586	1	-2,0586
2,19	+0,06	0,1314	1	0,1314
2,19	1,06	2,3214	10	23,2140
2,19	2,06	4,5114	18	81,2052
2,19	3,06	6,7014	6	40,2084
2,19	4,06	8,8914	1	8,8914
3,19	2,06	6,5714	7	45,9998
3,19	3,06	9,7614	12	117,1368
3,19	4,06	12,9514	6	77,7084
A reporter... ..			400	

		Report...	400	
4,19	3,06	12,8214			7	89,7498
4,19	4,06	17,0114			9	153,1026
4,19	5,06	21,2014			3	63,6042
4,19	6,06	25,3914			2	50,7828
5,19	4,06	21,0714			1	21,0714
5,19	5,06	26,2614			1	26,2614
5,19	6,06	31,4514			10	314,5140
6,19	6,06	37,5114			1	37,5114
8,19	8,06	66,0114			1	66,0114
					435	$\Sigma = 2190,3490$

$$E_{ab}^2 = \frac{1}{435} \Sigma = 5,0353$$

$$\sigma_a \cdot \sigma_b = 2,28 \times 2,33 = 5,3124$$

$$\text{Coefficient de corrélation : } \rho_{ab} = \frac{E_{ab}^2}{\sigma_a \sigma_b} = 0,948$$

$$\text{Coefficient de régression de } b \text{ en } d : \rho_{ab} \frac{\sigma_b}{\sigma_a} = \frac{5,0353}{5,2155} = 0,965$$

$$\text{Coefficient de régression de } d \text{ en } b : \rho_{ab} \frac{\sigma_a}{\sigma_b} = \frac{5,0353}{5,4390} = 0,925 = \frac{1}{1,080}$$

$$\text{Coefficient angulaire pour } \rho = 1 : \frac{\sigma_b}{\sigma_a} = \frac{2,33}{2,28} = 1,022$$

Calcul de l'écart rectangulaire E^2_{dg} et des coefficients ρ_{dg} et de régression de g en d et d en g .

$d - d_0$	$g - g_0$	$(d - d_0) (g - g_0)$	Fréquences	$F \times (d - d_0) (g - g_0)$
-4,81	-2,13	10,2453	3	30,7359
-3,81	-2,13	8,1153	11	89,2683
-2,81	-2,13	5,9853	24	143,6472
-2,81	-1,13	3,1753	25	79,3825
-2,81	-0,13	0,3653	1	0,3653
-1,81	-2,13	3,8553	1	3,8553
-1,81	-1,13	2,0453	64	130,8992
-1,81	-0,13	0,2353	10	2,3530
-1,81	+0,87	-1,5747	6	-9,4482
-0,81	-1,13	0,9153	27	24,7131
-0,81	-0,13	0,1053	42	4,4226
-0,81	+0,87	-0,7047	4	-2,8188
+0,19	-1,13	-0,2147	1	-0,2147
0,19	-0,13	-0,0247	45	-1,1115
0,19	+0,87	0,1653	18	2,9754
1,19	-0,13	-0,1547	19	-2,9393
1,19	0,87	1,0353	33	34,1649
1,19	1,87	2,2253	4	8,9012
2,19	-0,13	-0,2847	2	-0,5647
2,19	0,87	1,9053	26	49,5378
2,19	1,87	4,0953	9	36,8577
3,19	0,87	2,7753	4	11,1012
3,19	1,87	5,9653	17	101,4101
3,19	2,87	9,1553	4	36,6212
4,19	0,87	3,6453	2	7,2906
4,19	1,87	7,8353	13	101,8589
4,19	2,87	12,0253	6	72,1518
5,19	1,87	9,7053	4	38,8212
5,19	2,87	14,8953	5	74,4765
5,19	3,87	20,0853	3	60,2559
6,19	3,87	23,9553	1	23,9553
8,19	4,87	39,8853	1	39,8853
			435	$\Sigma = 1192,8102$

$$E^2_{dg} = \frac{1}{435} = 2,7421$$

$$\sigma_a \sigma_g = 2,28 \times 1,34 = 3,0552$$

$$\text{Coefficient de corrélation : } \rho_{dg} = \frac{E^2_{dg}}{\sigma_a \sigma_g} = \frac{2,7421}{3,0552} = 0,898$$

$$\text{Coefficient de régression de } g \text{ en } d : \rho_{dg} \frac{\sigma_g}{\sigma_a} = \frac{2,7421}{5,2155} = 0,526$$

$$\text{Coefficient de régression de } d \text{ en } g : \rho_{dg} \frac{\sigma_a}{\sigma_g} = \frac{2,7421}{1,7932} = 1,529 = \frac{1}{0,654}$$

$$\text{Coefficient angulaire pour } \rho = 1 : \frac{\sigma_g}{\sigma_a} = \frac{1,34}{2,28} = 0,588$$

Gisement d'Angleur.

[*Anthraconauta belgica* Auctorum.]Calcul de l'écart rectangulaire E_{da}^2 et des coefficients ρ_{da} et de régression de a en d et d en a .

$d - d_0$	$a - a_0$	$(d - d_0)(a - a_0)$	Fréquences	$F \times (d - d_0)(a - a_0)$
-4,11	-2,63	10,8093	14	151,3302
-2,11	-2,63	5,5493	26	144,2818
-2,11	-0,63	1,3293	46	61,1478
-0,11	-0,63	0,0693	35	2,4255
-0,11	+1,37	-0,1507	15	-2,2605
1,89	-0,63	-1,1907	4	-4,7628
1,89	+1,37	2,5893	41	106,1613
1,89	3,37	6,3693	4	25,4772
3,89	-0,63	-2,4507	1	-2,4507
3,89	1,37	5,3293	10	53,2930
3,89	3,37	13,1093	6	78,6558
3,89	5,37	20,8893	1	20,8893
5,89	1,37	8,0693	1	8,0693
5,89	5,37	31,6293	1	31,6293
7,89	3,37	26,5893	1	26,5893
9,89	5,37	53,1093	1	53,1093
9,89	9,37	92,6693	1	92,6693
11,89	5,37	63,8493	1	63,8493
			209	$\Sigma = 910,1037$

$$E_{da}^2 = 4,3546$$

$$\sigma_d \sigma_a = 2,64 \times 1,91 = 5,0424$$

$$\text{Coefficient de corrélation : } \rho_{da} = \frac{E_{da}^2}{\sigma_d \sigma_a} = \frac{4,3546}{5,0424} = 0,864$$

$$\text{Coefficient de régression de } a \text{ en } d : \rho_{da} \frac{\sigma_a}{\sigma_d} = \frac{4,3546}{6,9932} = 0,623$$

$$\text{Coefficient de régression de } d \text{ en } a : \rho_{da} \frac{\sigma_d}{\sigma_a} = \frac{4,3546}{3,6631} = 1,189 = \frac{1}{0,841}$$

$$\text{Coefficient angulaire pour } \rho = 1 : \frac{\sigma_a}{\sigma_d} = 0,723$$

Calcul de l'écart rectangulaire E_{ab}^2 et des coefficients ρ_{ab} et de régression de b en d et d en b .

$d - d_0$	$b - b_0$	$(d - d_0)(b - b_0)$	Fréquences	$F \times (d - d_0)(b - b_0)$
-4,11	-4,45	18,2895	7	128,0265
-4,11	-2,45	10,0695	7	70,4865
-2,11	-4,45	9,3895	3	28,1685
-2,11	-2,45	5,1695	59	305,0005
-2,11	-0,45	0,9495	10	9,4950
-0,11	-2,45	0,2695	2	0,5390
-0,11	-0,45	0,0495	38	1,8810
-0,11	1,55	-0,1705	10	-1,7050
1,89	-0,45	-0,8505	5	-4,2525
1,89	1,55	2,9295	28	82,0260
1,89	3,55	6,7095	16	107,3520
3,89	1,55	6,0295	2	12,0590
3,89	3,55	13,8095	9	124,2855
3,89	5,55	21,5895	7	151,1265
5,89	5,55	32,6895	2	65,3790
7,89	5,55	43,7895	1	43,7895
9,89	7,55	74,6695	1	74,6695
9,89	11,55	114,2295	1	114,2295
11,89	9,55	113,5495	1	113,5495
			209	$\Sigma = 1426,1055$

$$E_{ab}^2 = 6,8235$$

$$\sigma_d \sigma_b = 2,64 \times 2,78 = 7,3392$$

$$\text{Coefficient de corrélation : } \rho_{ab} = \frac{6,8235}{7,3392} = 0,930$$

$$\text{Coefficient de régression de } b \text{ en } d : \rho_{ab} \frac{\sigma_b}{\sigma_d} = \frac{6,8235}{6,9932} = 0,976$$

$$\text{Coefficient de régression de } d \text{ en } b : \rho_{ab} \frac{\sigma_d}{\sigma_b} = \frac{6,8235}{7,7372} = 0,882 = \frac{1}{1,134}$$

$$\text{Coefficient angulaire pour } \rho = 1 : \frac{\sigma_d}{\sigma_b} = \frac{2,78}{2,64} = 1,053$$

ÉTUDE BIOMÉTRIQUE DES ANTHRACONAUTA

Calcul de l'écart rectangulaire E_{dg}^2 et des coefficients ρ_{dg} et de régression de g en d et d en g .

$d - d_0$	$g - g_0$	$(d - d_0)(g - g_0)$	Fréquences	$F \times (d - d_0)(g - g_0)$
-4,08	-2,39	9,7512	11	107,2632
-4,08	-1,39	5,6712	4	22,6848
-2,08	-2,39	4,9712	10	49,7120
-2,08	-1,39	2,8912	41	118,5392
-2,08	-0,39	0,8112	14	11,3568
-2,08	0,61	-1,2688	3	-3,8064
-2,08	+1,61	-3,3488	2	-6,6976
-0,08	-1,39	0,1112	5	0,5560
-0,08	-0,39	0,0312	25	0,7800
-0,08	0,61	-0,5488	14	-0,6832
-0,08	1,61	-0,1288	2	-0,2576
1,92	-1,39	-2,6688	1	-2,6688
1,92	-0,39	-0,7488	4	-2,9952
1,92	0,61	1,1712	17	19,9104
1,92	1,61	3,0912	20	61,8240
1,92	2,61	5,0112	5	25,0560
1,92	3,61	6,9312	1	6,9312
3,92	-0,39	-1,5288	1	-1,5288
3,92	0,61	2,3912	4	9,5648
3,92	1,61	6,3112	4	25,2448
3,92	2,61	10,2312	6	61,3872
3,92	3,61	14,1512	2	28,3024
5,92	3,61	21,3712	1	21,3712
5,92	4,61	27,2912	1	27,2912
7,92	3,61	28,5912	1	28,5912
9,92	5,61	55,6512	1	55,6512
9,92	6,61	65,5712	1	65,5712
11,92	6,61	78,7912	1	78,7912
			202	$\Sigma = 807,7424$

$$E_{dg}^2 = 3,9987$$

$$\sigma_d \sigma_g = 2,69 \times 1,59 = 4,2771$$

$$\text{Coefficient de corrélation : } \rho_{dg} = \frac{3,9987}{4,2771} = 0,935$$

$$\text{Coefficient de régression de } g \text{ en } d : \rho_{dg} \frac{\sigma_g}{\sigma_d} = \frac{3,9987}{7,2316} = 0,553$$

$$\text{Coefficient de régression de } d \text{ en } g : \rho_{dg} \frac{\sigma_d}{\sigma_g} = \frac{3,9987}{2,5312} = 1,580 = \frac{1}{0,633}$$

$$\text{Coefficient angulaire pour } \rho = 1 : \frac{\sigma_g}{\sigma_d} = 0,591$$

Gisement de Stud.

[*Anthraconauta belgica* Auctorum.]

Calcul de l'écart rectangulaire E_{da}^2 et des coefficients ρ_{da} et de régression de a en d et d en a .

$\bar{d} - d_0$	$a - a_0$	$(\bar{d} - d_0)(a - a_0)$	Fréquences	$F \times (\bar{d} - d_0)(a - a_0)$
-4,37	-2,36	10,3132	6	61,8792
-3,37	-2,36	7,9532	11	87,4852
-3,37	-1,36	4,5832	2	9,1664
-2,37	-2,36	5,5932	5	27,9660
-2,37	-1,36	3,2232	12	38,6784
-2,37	-0,36	0,8532	2	1,7064
-1,37	-2,36	3,2332	2	6,4664
-1,37	-1,36	1,8632	9	16,7688
-1,37	-0,36	0,4932	3	1,4796
-1,37	+0,64	-0,8768	1	-0,8768
-0,37	-1,36	0,5032	4	2,0128
-0,37	-0,36	0,1332	10	1,3320
-0,37	+0,64	-0,2368	8	-1,8944
-0,37	1,64	-0,6068	2	-1,2136
+0,63	-1,36	-0,8568	1	-0,8568
+0,63	-0,36	-0,2268	14	-3,1752
0,63	0,64	0,4032	8	3,2256
0,63	1,64	1,0332	6	6,1992
1,63	-0,36	-0,5868	1	-0,5868
1,63	0,64	1,0432	4	4,1728
1,63	1,64	2,6732	9	24,0588
2,63	0,64	1,6832	2	3,3664
2,63	1,64	4,3132	5	21,5660
2,63	2,64	6,9432	4	27,7728
3,63	1,64	5,9532	7	41,6724
3,63	2,64	9,5832	3	28,7496
3,63	3,64	13,2132	2	26,4264
4,63	2,64	12,2232	2	24,4464
4,63	3,64	16,8532	1	16,8532
5,63	2,64	14,8632	2	29,7264
5,63	3,64	20,4932	1	40,9864
			149	$\Sigma = 545,5600$

$$E_{da}^2 = 3,6615$$

$$\sigma_d \sigma_a = 2,37 \times 1,66 = 3,9342$$

$$\text{Coefficient de corrélation : } \rho_{da} = \frac{E_{da}^2}{\sigma_d \sigma_a} = \frac{3,6614}{3,9342} = 0,931$$

$$\text{Coefficient de régression de } a \text{ en } d : \rho_{da} \frac{\sigma_a}{\sigma_d} = \frac{3,6215}{5,6428} = 0,649$$

$$\text{Coefficient de régression de } d \text{ en } a : \rho_{da} \frac{\sigma_d}{\sigma_a} = \frac{3,6615}{2,7536} = 1,330 = \frac{1}{0,752}$$

$$\text{Coefficient angulaire pour } \rho = 1 : \frac{\sigma_a}{\sigma_d} = \frac{1,66}{2,37} = 0,700$$

Calcul de l'écart rectangulaire E_{ab}^2 et des coefficients ρ_{ab} et de régression de b en d et d en b .

$d - d_0$	$g - g_0$	$(d - d_0)(g - g_0)$	Fréquences	$F \times (d - d_0)(g - g_0)$
-4,37	-4,56	19,9272	2	39,8544
-4,37	-3,56	15,5572	3	46,6716
-4,37	-2,56	11,1872	1	11,1872
-3,37	-4,56	15,3672	1	15,3672
-3,37	-3,56	11,9972	4	47,9888
-3,37	-2,56	8,6272	8	69,0176
-2,37	-3,56	8,4372	2	16,8744
-2,37	-2,56	6,0672	10	60,6720
-2,37	-1,56	3,6972	6	22,1832
-2,37	-0,56	1,3272	1	1,3272
-1,37	-3,56	4,8772	1	4,8772
-1,37	-2,56	3,5072	3	10,5216
-1,37	-1,56	2,1372	9	19,2348
-1,37	-0,56	0,7672	1	0,7672
-1,37	+0,44	-0,6028	1	-0,6028
-0,37	-1,56	0,5772	4	2,3088
-0,37	-0,56	0,2072	12	2,4864
-0,37	+0,44	-0,1628	7	-1,1396
-0,37	+1,44	-0,5328	1	-0,5328
+0,63	-1,56	-0,9828	1	-0,9828
+0,63	-0,56	-0,3528	3	-1,0584
+0,63	0,44	0,2772	17	4,7124
0,64	1,44	0,9072	7	6,3504
0,63	2,44	1,5372	1	1,5372
1,63	0,44	0,7172	5	3,5860
1,63	1,44	2,3472	7	16,4304
1,63	2,44	3,9772	2	7,9544
2,63	1,44	3,7872	3	11,3616
2,63	2,44	6,4172	2	12,8344
2,63	3,44	9,0472	5	45,2360
2,63	4,44	11,6772	1	11,6772
3,63	1,44	5,2272	1	5,2272
3,63	2,44	8,8572	2	17,7144

A reporter... .. 134

		Report...	134	
3,63	3,44	12,4872		4	49,9488
3,63	4,44	16,1171		3	48,3516
3,63	5,44	19,7472		2	39,4944
4,63	3,44	15,9272		1	15,9272
4,63	4,44	20,5572		2	41,1144
5,63	4,44	24,9972		1	24,9972
5,63	5,44	30,6272		1	30,6272
5,63	7,44	41,8872		1	41,8872
				149	$\Sigma = 803,9928$

$$E_{ab}^2 = 5,3959$$

$$\sigma_a \sigma_b = 2,37 \times 2,43 = 5,759$$

$$\text{Coefficient de corrélation : } \rho_{ab} = \frac{E_{ab}^2}{\sigma_a \sigma_b} = \frac{5,3959}{5,7591} = 0,937$$

$$\text{Coefficient de régression } b \text{ en } d : \rho_{ab} \frac{\sigma_b}{\sigma_a} = \frac{5,3959}{5,6428} = 0,956$$

$$\text{Coefficient de régression } d \text{ en } b : \rho_{ab} \frac{\sigma_a}{\sigma_b} = \frac{5,3959}{5,9269} = 0,910 = \frac{1}{1,098}$$

$$\text{Coefficient angulaire pour } \rho : = 1 \frac{\sigma_b}{\sigma_a} = 1,025$$

Calcul de l'écart rectangulaire E_{dg}^2 et des coefficients ρ_{dg} et de régression de g en d et d en g .

$d - d_0$	$b - b_0$	$(d - d_0)(b - b_0)$	Fréquences	$F \times (d - d_0)(a - a_0)$
-4,37	-1,44	6,2928	1	6,2928
-4,37	-2,44	10,6628	5	54,3140
-3,37	-2,44	8,2228	5	41,1140
-3,37	-1,44	4,8528	7	33,9696
-3,37	-0,44	1,4828	1	1,4828
-2,37	-2,44	5,7828	2	11,5656
-2,37	-1,44	3,4128	13	44,3664
-2,37	-0,44	1,0428	4	4,1712
-1,37	-2,44	3,3428	1	3,3428
-1,37	-1,44	1,9728	7	13,8096
-1,37	-0,44	0,6028	6	3,6168
-1,37	+0,56	-0,7672	1	-0,7672
-0,37	-1,44	0,5328	1	0,5328
-0,37	-0,44	0,1628	13	2,1164
-0,37	+0,56	-0,2072	8	-1,6576
-0,37	+1,56	-0,5772	2	-1,1544
+0,63	-1,44	-0,9072	1	-0,9072
+0,63	-0,44	-0,2772	6	-1,6632
0,63	0,56	0,3528	20	7,0560
0,63	1,56	0,9828	2	1,9656
1,63	0,56	0,9128	11	10,0408
1,64	1,56	2,5428	3	7,6284
2,63	0,56	1,4728	3	4,4184
2,63	1,56	4,1028	5	20,5140
2,63	2,56	6,7328	3	20,1984
3,63	0,56	2,0328	2	4,0656
3,63	1,56	5,6628	5	28,3140
3,63	2,56	9,2928	4	37,1712
3,63	3,56	12,9228	1	12,9228
4,63	1,56	7,2228	2	14,4456
4,63	3,56	16,4828	1	16,4828
5,63	2,56	14,4128	2	28,8256
5,63	3,56	20,0428	1	20,0428
			149	$\Sigma = 447,6372$

$$E_{dg}^2 = 3,0043$$

$$\text{Coefficient de corrélation : } \rho_{dg} = \frac{E_{dg}^2}{\sigma_d \sigma_g} = \frac{3,0043}{3,3891} = 0,882$$

$$\text{Coefficient de régression de } g \text{ en } d : \rho_{dg} \frac{\sigma_g}{\sigma_d} = \frac{3,0043}{5,6428} = 0,532$$

$$\text{Coefficient de régression de } d \text{ en } g : \rho_{dg} \frac{\sigma_d}{\sigma_g} = \frac{3,0043}{2,0582} = 1,460 = \frac{1}{0,685}$$

$$\text{Coefficient angulaire pour } \rho = 1 : \frac{\sigma_g}{\sigma_d} = \frac{1,43}{2,37} = 0,603$$

CONCLUSIONS.

Il a été réuni, dans le tableau ci-dessous, les valeurs des coefficients de corrélation ρ_{da} , ρ_{ab} , ρ_{dg} , pour chacune des trois populations étudiées.

Origine	Coefficients		
	ρ_{da}	ρ_{ab}	ρ_{dg}
A. — Wérister	0,940	0,948	0,898
A. — Angleur	0,864	0,930	0,935
A. — Stud	0,931	0,937	0,886

Ensuite il a été calculé les écarts quadratiques dont l'utilisation serait légitime si les distributions étaient régies par une loi de BRAVAIS. Rappelons que ces écarts quadratiques, calculés au moyen de la formule

$$\sigma = \frac{1 - \rho^2}{\sqrt{N}},$$

nous fixent l'ordre de grandeur des écarts à partir de la valeur théorique exacte de ρ .

Origine	Ecart quadratiques		
	$\sigma_{\rho_{da}}$	$\sigma_{\rho_{ab}}$	$\sigma_{\rho_{dg}}$
A. — Wérister	0,0058	0,0049	0,0093
A. — Angleur	0,0175	0,0093	0,0087
A. — Stud	0,0109	0,0099	0,0176

Remarque. — Dans la présente étude interviennent simultanément plusieurs caractères : d , a , b , g . PEARSON ⁽⁶²⁾ a fait observer qu'une telle étude nécessite, en général, l'introduction d'autres coefficients que les coefficients de corrélation relatifs aux combinaisons des divers caractères, pris deux à deux [G. U. YULE et M. KENDALL ⁽⁶³⁾].

Mais la quasi-linéarité des relations entre les couples de variables correspondantes suffit ici pour établir les conclusions relatives à la stabilité de forme géométrique des coquilles, qu'on trouvera plus loin.

⁽⁶³⁾ UDN YULE, G., et KENDALL, M. G., 1945, pp. 261-286.

E. — ÉTUDE DES INDICES CARACTÉRISTIQUES.

(Rapport entre mesures des caractères des individus. Distribution de fréquences.)

MÉTHODE. — Selon les méthodes d'établissement de relevés statistiques portant sur des variables continues, exposés au chapitre II, il a été procédé à un relevé direct des distributions de fréquences des indices caractéristiques.

Ces indices caractéristiques, étudiés dans les trois cas pour A. — Angleur, A. — Stud et A. — Wérister, sont les rapports

$$100 \times \frac{a}{b}, \quad 100 \times \frac{d}{a}, \quad 100 \times \frac{b}{g},$$

obtenus en effectuant les quotients de deux des grandeurs caractéristiques mesurées sur les coquilles.

Désignant par x , de façon générale, ces indices, il a été compté dans chaque intervalle partiel de longueur h convenablement choisie, centré sur les points x_i , les fréquences absolues $\varphi(x_i)h$; ce qui a permis de déterminer les valeurs expérimentales de la fonction à paliers $\varphi(x_i)$.

Les moyennes et dispersions permettent de déterminer la forme analytique de la loi de GAUSS :

$$p(x) = \frac{N}{\sigma \sqrt{2\pi}} e^{-\frac{(x-A)^2}{2\sigma^2}},$$

qui, en première et suffisante approximation, réalise l'ajustement de la distribution des fréquences absolues.

Pour chacune des trois populations étudiées et pour chacun des indices caractéristiques, les résultats (relevés bruts), les calculs et ajustements sont présentés ci-dessous de façon uniforme.

N est le nombre de coquilles mesurées, n le nombre d'intervalles partiels et h la longueur de ces intervalles, A et σ la moyenne et la dispersion. Le coefficient de concentration $\frac{\sigma}{A}$ donne une idée de l'importance relative de la dispersion par rapport à la moyenne.

$\frac{N}{\sigma \sqrt{2\pi}}$ est l'ordonnée maximum de la courbe de GAUSS qui réalise l'ajustement des distributions représentées dans les figures des planches X à XIV sous la forme de polygones de rectangles. Ce maximum théorique est atteint pour une valeur de x égale à A .

Dans les tableaux rassemblant les relevés et les résultats des ajustements, les titres des colonnes ont la signification suivante :

- 1) x_i : centres des intervalles partiels de longueur h ;
- 2) X_i : valeurs correspondantes de la variable auxiliaire X définie par $x = hX + x_0$;

- 3) $\varphi(x_i)h$: fréquences absolues;
- 4) $\varphi(x_i)$: ordonnées de la fonction à paliers;
- 5) $x_i - A$: écarts de la moyenne;
- 6) $t = \frac{x_i - A}{\sigma}$: seconde variable auxiliaire utilisée pour l'emploi des tables

donnant les valeurs de la fonction $y = \frac{1}{\sqrt{\pi}} e^{-\frac{t^2}{2}}$;

7) $\frac{1}{\sqrt{2\pi}} e^{-\frac{t^2}{2}}$: valeurs tirées d'une table;

8) $\frac{N}{\sigma\sqrt{2\pi}} e^{-\frac{t^2}{2}}$: valeurs ajustées de la fonction $\varphi(x_i)$ ordonnées de la

courbe de GAUSS, correspondant aux abscisses x_i .

§ 5. DISTRIBUTION DE FRÉQUENCES.

Gisement de Wérister.

[*Anthraconauta minima* HIND (non LUDWIG), espèce de référence.]

1. Indice $100 \times \frac{a}{b}$.

N = 433	A = 71,94	$\frac{N}{\sigma} = 40,243$
n = 16	$\sigma = 10,76$	
h = 5	$\frac{\sigma}{A} = 0,150$	$\frac{N}{\sigma\sqrt{2}} = 16,054$

$100 \frac{a}{b} = x_i$	X_i	$\varphi(x_i)h$	$\varphi(x_i)$	$x_i - A$	$t = \frac{x_i - A}{\sigma}$	$\frac{1}{\sqrt{2\pi}} e^{-\frac{t^2}{2}}$	$\frac{N}{\sigma\sqrt{2\pi}} e^{-\frac{t^2}{2}}$
50	-4	5	1	-21,94	-2,04	0,04999	2,012
55	-3	15	3	-16,94	-1,57	0,11650	4,688
60	-2	51	10,2	-11,94	-1,11	0,21548	8,671
65	-1	94	18,8	-6,94	-0,64	0,32483	13,072
70	0	96	19,2	-1,94	-0,18	0,39222	15,784
75	1	64	12,8	3,06	0,28	0,38332	15,426
80	2	36	7,2	8,06	0,75	0,30097	12,112
85	3	34	6,8	13,06	1,21	0,19191	7,723
90	4	19	3,8	18,06	1,68	0,09742	3,920
95	5	8	1,6	23,06	2,14	0,04058	1,633
100	6	8	1,6	28,06	2,61	0,01326	0,533
105	7	1	0,2	33,06	3,07	0,00362	0,146
110	8	—	—	—	—	—	—
115	9	1	0,2	43,06	4,00	0,00013	0,005
120	10	—	—	—	—	—	—
125	11	1	0,2	53,06	4,93	—	—

Ecart probable : $0,6745 \times 10,76 = 7,258$.

2. Indice $100 \times \frac{\bar{d}}{a}$.

$$N = 433$$

$$A = 135,77$$

$$\frac{N}{\sigma} = 32,322$$

$$n = 20$$

$$\sigma = 13,38$$

$$h = 5$$

$$\frac{\sigma}{A} = 0,099$$

$$\frac{N}{\sigma \sqrt{2\pi}} = 12,910$$

$100 \frac{\bar{d}}{a} = x_i$	X_i	$\varphi(x_i) h$	$\varphi(x_i)$	$x_i - A$	$t = \frac{x_i - A}{\sigma}$	$\frac{1}{\sqrt{2\pi}} e^{-\frac{t^2}{2}}$	$\frac{N}{\sigma \sqrt{2\pi}} e^{-\frac{t^2}{2}}$
90	-10	1	0,2	-45,77	-3,42	0,00116	0,038
95	-9	0	0	-40,77	-3,05	0,00385	0,125
100	-8	1	0,2	-35,77	-2,67	0,01137	0,368
105	-7	4	0,8	-30,77	-2,30	0,02833	0,917
110	-6	6	1,2	-25,77	-1,93	0,06213	2,011
115	-5	22	4,4	-20,77	-1,55	0,12022	3,891
120	-4	34	6,8	-15,77	-1,18	0,19892	6,437
125	-3	40	8,0	-10,77	-0,80	0,28969	9,375
130	-2	50	10,0	-5,77	-0,43	0,36341	11,761
135	-1	53	10,6	-0,77	-0,06	0,39775	12,872
140	0	78	15,6	4,23	0,32	0,37877	12,258
145	1	47	9,4	9,23	0,69	0,31015	10,037
150	2	58	11,6	14,23	1,06	0,22750	7,362
155	3	16	3,2	19,23	1,44	0,14165	4,584
160	4	7	1,4	24,23	1,81	0,07762	2,512
165	5	6	1,2	29,23	2,18	0,03717	1,203
170	6	6	1,2	34,23	2,56	0,01516	0,491
175	7	3	0,6	39,23	2,93	0,00549	0,178
180	8	0	0	44,23	3,31	0,00167	0,054
185	9	1	0,2	49,23	3,68	0,00046	0,015

Ecart probable : $0,6745 \times 13,38 = 9,025$.

3. Indice $100 \times \frac{a}{b}$.

$N = 425$

$A = 179,52$

$\frac{N}{\sigma} = 20,722$

$n = 14$

$\sigma = 20,51$

$h = 10$

$\frac{\sigma}{A} = 0,114$

$\frac{N}{\sigma \sqrt{2\pi}} = 8,267$

$100 \frac{b}{g} = x_i$	X_i	$\varphi(x_i) h$	$\varphi(x_i)$	$x_i - A$	$t = \frac{x_i - A}{\sigma}$	$\frac{1}{\sqrt{2\pi}} e^{-\frac{t^2}{2}}$	$\frac{N}{\sigma \sqrt{2\pi}} e^{-\frac{t^2}{2}}$
115	-6	7	0,7	-64,52	-3,15	0,00283	0,059
125	-5	3	0,3	-54,52	-2,66	0,01168	0,042
135	-4	4	0,4	-44,52	-2,17	0,03802	0,788
145	-3	15	1,5	-34,52	-1,68	0,09742	1,999
155	-2	25	2,5	-24,52	-1,20	0,19419	4,024
165	-1	68	6,8	-14,52	-0,71	0,31000	6,424
175	0	93	9,3	- 4,52	-0,22	0,38911	8,063
185	1	91	9,1	+ 5,48	0,27	0,38429	7,963
195	2	62	6,2	15,48	0,75	0,30097	6,237
205	3	29	2,9	25,48	1,24	0,18506	3,835
215	4	17	1,7	35,48	1,73	0,08952	1,855
225	5	9	0,9	45,48	2,22	0,03404	0,705
235	6	1	0,1	55,48	2,71	0,01017	0,211
245	7	1	0,1	65,48	3,19	0,00229	0,047

Ecart probable : $0,6745 \times 20,51 = 13,833$.

Gisement d'Angleur.

[*Anthraconauta belgica* Auctorum.]1. Indice $100 \times \frac{a}{b}$.

$N = 209$

$A = 72,92$

$\frac{N}{\sigma} = 14,8227$

$h = 10$

$\sigma = 14,10$

$n = 10$

$\frac{\sigma}{A} = 0,1934$

$\frac{N}{\sigma \sqrt{2\pi}} = 5,913$

$100 \frac{a}{b} = x_i$	X_i	$\varphi(x_i) h$	$\varphi(x_i)$	$x_i - A$	$t = \frac{x_i - A}{\sigma}$	$\frac{1}{\sqrt{2\pi}} e^{-\frac{t^2}{2}}$	$\frac{N}{\sigma \sqrt{2\pi}} e^{-\frac{t^2}{2}}$
50	-2	15	1,5	-22,92	-1,6	0,11092	1,64
60	-1	51	5,1	-12,92	-0,9	0,26609	3,94
70	0	64	6,4	-2,92	-0,2	0,39104	5,80
80	1	40	4,0	7,08	0,5	0,35207	5,22
90	2	27	2,7	17,08	1,2	0,19419	2,88
100	3	7	0,7	27,08	1,9	0,06562	0,97
110	4	2	0,2	37,08	2,6	0,01358	0,20
120	5	0	0	47,08	3,3	0,00172	0,02
130	6	2	0,2	57,08	4	0,00013	—
140	7	1	0,1	67,08	4,7	0,00001	—

Ecart probable : $0,6745 \times 14,10 = 9,51$.2. Indice $100 \times \frac{d}{a}$.

$N = 209$

$A = 131,26$

$N = 9,9335$

$h = 10$

$\sigma = 21,04$

$n = 13$

$\frac{\sigma}{A} = 0,1603$

$\frac{N}{\sigma \sqrt{2\pi}} = 3,9629$

$100 \frac{d}{a} = x_i$	X_i	$\varphi(x_i) h$	$\varphi(x_i)$	$x_i - A$	$t = \frac{x_i - A}{\sigma}$	$\frac{1}{\sqrt{2\pi}} e^{-\frac{t^2}{2}}$	$\frac{N}{\sigma \sqrt{2\pi}} e^{-\frac{t^2}{2}}$
77	-5	1	0,1	-54,26	-2,6	0,01358	0,13
87	-4	1	0,1	-44,26	-2,1	0,04398	0,44
97	-3	5	0,5	-34,26	-1,6	0,11092	1,10
107	-2	23	2,3	-24,26	-1,15	0,20602	2,05
117	-1	40	4,0	-14,26	-0,7	0,31225	3,10
127	0	48	4,8	-4,26	-0,2	0,39104	3,88
137	1	44	4,4	6,74	0,3	0,38139	3,79
147	2	20	2,0	16,74	0,8	0,28969	2,88
157	3	12	1,2	26,74	1,3	0,17137	1,70
167	4	5	0,5	36,74	1,7	0,09405	0,93
177	5	3	0,3	46,74	2,2	0,03547	0,35
187	6	5	0,5	56,74	2,7	0,01042	0,10
197	7	2	0,2	66,74	3,2	0,00238	0,02

Ecart probable : $0,6745 \times 21,04 = 14,19$.

3. Indice $100 \times \frac{b}{g}$.

$N = 201$

$A = 165,27$

$\frac{N}{\sigma} = 7,7308$

$h = 10$

$\sigma = 26,0$

$n = 18$

$\frac{\sigma}{A} = 0,1573$

$\frac{N}{\sigma \sqrt{2\pi}} = 3,0841$

$100 \frac{b}{g} = x_i$	X_i	$\varphi(x_i) h$	$\varphi(x_i)$	$x - A$	$t = \frac{x_i - A}{\sigma}$	$\frac{1}{\sqrt{2\pi}} e^{-\frac{t^2}{2}}$	$\frac{N}{\sigma \sqrt{2\pi}} e^{-\frac{t^2}{2}}$
80	-8	1	0,1	-85,27	-3,3	0,00172	0,01
90	-7	2	0,2	-75,27	-2,9	0,00595	0,05
100	-6	1	0,1	-65,27	-2,5	0,01753	0,14
110	-5	3	0,3	-55,27	-2,1	0,04398	0,34
120	-4	5	0,5	-45,27	-1,7	0,09405	0,73
130	-3	9	0,9	-35,27	-1,35	0,16055	1,24
140	-2	13	1,3	-25,27	-1,0	0,24191	1,87
150	-1	27	2,7	-15,27	-0,6	0,33322	2,58
160	0	42	4,2	- 5,27	-0,2	0,39104	3,02
170	1	37	3,7	4,73	+0,2	0,39104	3,02
180	2	18	1,8	14,73	+0,55	0,34264	2,65
190	3	19	1,9	24,73	0,95	0,25403	1,96
200	4	11	1,1	34,73	1,3	0,17137	1,32
210	5	8	0,8	44,73	1,7	0,09405	0,73
220	6	2	0,2	54,73	2,1	0,04398	0,34
230	7	2	0,2	64,73	2,5	0,01753	0,14
240	8	—	—	74,73	2,9	0,00595	0,05
250	9	1	0,1	84,73	3,3	0,00172	0,01

Fcart probable : $0,6745 \times 26 = 17,54$.

Gisement de Stud.

[*Anthraconauta belgica* Auctorum.]1. Indice $100 \times \frac{a}{b}$.

$N = 150$

$A = 72,10$

$\frac{N}{\sigma} = 11,96$

$n = 10$

$\sigma = 12,54$

$h = 7$

$\frac{\sigma}{A} = 0,174$

$\frac{N}{\sigma \sqrt{2\pi}} = 4,7720$

$100 \frac{a}{b} = x_i$	X_i	$\varphi(x_i) h$	$\varphi(x_i)$	$x_i - A$	$t = \frac{x_i - A}{\sigma}$	$\frac{1}{\sqrt{2\pi}} e^{-\frac{t^2}{2}}$	$\frac{N}{\sigma \sqrt{2\pi}} e^{-\frac{t^2}{2}}$
42	-5	1	0,14	-30,1	-2,40	0,02239	0,268
49	-4	5	0,71	-23,1	-1,84	0,07362	0,880
56	-3	17	2,43	-16,1	-1,28	0,17593	2,104
63	-2	31	4,43	-9,1	-0,73	0,30548	3,554
70	-1	29	4,14	-2,1	-0,17	0,39281	4,698
77	0	33	4,71	+ 4,9	0,39	0,36958	4,420
84	1	16	2,29	11,9	0,95	0,25403	3,038
91	2	11	1,57	18,9	1,51	0,12766	1,527
98	3	4	0,57	25,9	2,07	0,04098	0,490
105	4	3	0,43	32,9	2,62	0,01295	0,155

Ecart probable : $0,6745 \times 12,54 = 8,46$.2. Indice $100 \times \frac{d}{a}$.

$N = 150$

$A = 136,98$

$\frac{N}{\sigma} = 8,681$

$n = 15$

$\sigma = 17,28$

$h = 7$

$\frac{\sigma}{A} = 0,126$

$\frac{N}{\sigma \sqrt{2\pi}} = 3,4630$

$100 \frac{d}{a} = x_i$	X_i	$\varphi(x_i) h$	$\varphi(x_i)$	$x_i - A$	$t = \frac{x_i - A}{\sigma}$	$\frac{1}{\sqrt{2\pi}} e^{-\frac{t^2}{2}}$	$\frac{N}{\sigma \sqrt{2\pi}} e^{-\frac{t^2}{2}}$
101	-5	2	0,29	-35,98	-2,08	0,04598	0,399
108	-4	3	0,43	-28,98	-1,68	0,09737	0,845
115	-3	18	2,57	-21,98	-1,27	0,17882	1,547
122	-2	17	2,43	-14,98	-0,87	0,27317	2,371
129	-1	27	3,86	-7,98	-0,46	0,35855	3,113
136	0	20	2,86	-0,98	-0,06	0,39775	3,453
143	1	19	2,71	6,02	+0,35	0,37483	3,254
150	2	20	2,86	13,02	0,75	0,30097	2,613
157	3	12	1,71	20,02	1,16	0,22731	1,973
164	4	5	0,71	27,02	1,56	0,11836	1,027
171	5	2	0,29	34,02	1,97	0,05748	0,499
178	6	4	0,57	41,02	2,37	0,02417	0,210
185	7	—	—	48,02	2,78	0,00842	0,073
192	8	—	—	55,02	3,18	0,00256	0,022
199	9	1	0,14	62,02	3,59	0,00064	0,006

Ecart probable : $0,6754 \times 17,28 = 11,55$.

3. Indice $100 \times \frac{b}{g}$.

$$N = 146$$

$$A = 170,95$$

$$\frac{N}{\sigma} = 6,55$$

$$n = 13$$

$$\sigma = 22,24$$

$$h = 10$$

$$\frac{\sigma}{A} = 0,130$$

$$\frac{N}{\sigma \sqrt{2\pi}} = 2,2189$$

$100 \frac{b}{g} = x_i$	X_i	$\varphi(x_i) h$	$\varphi(x_i)$	$x_i - A$	$t = \frac{x_i - A}{\sigma}$	$\frac{1}{\sqrt{2\pi}} e^{-\frac{t^2}{2}}$	$\frac{N}{\sigma \sqrt{2\pi}} e^{-\frac{t^2}{2}}$
108	-6	3	0,3	-62,95	-2,83	0,00733	0,048
118	-5	1	0,1	-52,95	-2,38	0,02358	0,155
128	-4	1	0,1	-42,95	-1,93	0,06213	0,408
138	-3	1	0,1	-32,95	-1,48	0,13356	0,876
148	-2	23	2,3	-22,95	-1,03	0,23473	1,540
158	-1	28	2,8	-12,95	-0,58	0,33699	2,211
168	0	21	2,1	- 2,95	-0,13	0,39518	2,592
178	1	30	3,0	+ 7,05	+0,32	0,37877	2,485
188	2	19	1,9	+17,05	0,77	0,28292	1,856
198	3	7	0,7	+27,05	1,22	0,18963	1,244
208	4	5	0,5	+37,05	1,67	0,09911	0,650
218	5	4	0,4	+47,05	2,12	0,04228	0,277
228	6	3	0,3	+57,05	2,57	0,01477	0,097

Ecart probable : $0,6745 \times 22,24 = 15,00$.

§ 6. ÉTUDE DES ANGLES CARACTÉRISTIQUES.

Considérant toujours la mesure de la grandeur d comme une bonne estimation de l'âge des coquilles, il a été dressé des tables de répartition des couples de variables statistiques (d, α) et (d, β) . Ces tables ont révélé une absence de corrélation entre les caractères angulaires α , β et d aisément explicable et dont la signification sera dégagée plus loin. Par contre, les distributions de la fréquence des variables α et β sont très révélatrices; elles montrent une concentration nette des mesures autour de certaines valeurs (voir Pls. XV à XVIII et tableaux pp. 84 à 86).

Gisement de Wérister.

[*Anthraconauta minima* HIND (non LUDWIG), espèce de référence.]1. Angle α .

$$N = 421$$

$$A = 39^{\circ},66$$

$$\frac{N}{\sigma} = 67,252$$

$$n = 9$$

$$\sigma = 6^{\circ},26$$

$$h = 5$$

$$\frac{\sigma}{A} = 0,150$$

$$\frac{N}{\sigma \sqrt{2\pi}} = 26,830$$

$\alpha = \omega_i$	X_i	$\varphi(\omega_i) h$	$\varphi(\omega_i)$	$\omega_i - A$	$t = \frac{\omega_i - A}{\sigma}$	$\frac{1}{\sqrt{2\pi}} e^{-\frac{t^2}{2}}$	$\frac{N}{\sigma \sqrt{2\pi}} e^{-\frac{t^2}{2}}$
25	-3	3	0,6	-14,66	-2,34	0,02596	1,745
30	-2	46	9,2	-9,66	-1,54	0,12208	8,210
35	-1	118	23,6	-4,66	-0,74	0,30323	20,393
40	0	133	26,6	0,34	+0,05	0,39794	26,762
45	1	66	13,2	5,34	0,85	0,27789	18,689
50	2	43	8,6	10,34	1,65	0,10249	6,893
55	3	10	2,0	15,34	2,45	0,01996	1,342
60	4	1	0,2	20,34	3,25	0,00205	0,138
65	5	1	0,2	25,34	4,05	0,00011	0,007

Ecart probable : $0,6745 \times 6^{\circ},26 = 4^{\circ},222$.2. Angle β .

$$N = 423$$

$$A = 107^{\circ},91$$

$$\frac{N}{\sigma} = 45,582$$

$$n = 11$$

$$\sigma = 9^{\circ},28$$

$$h = 5$$

$$\frac{\sigma}{A} = 0,073$$

$$\frac{N}{\sigma \sqrt{2\pi}} = 18,184$$

$\beta = \omega_i$	X_i	$\varphi(\omega_i) h$	$\varphi(\omega_i)$	$\omega_i - A$	$t = \frac{\omega_i - A}{\sigma}$	$\frac{1}{\sqrt{2\pi}} e^{-\frac{t^2}{2}}$	$\frac{N}{\sigma \sqrt{2\pi}} e^{-\frac{t^2}{2}}$
100	-6	2	0,4	-27,91	-3,00	0,00443	0,202
105	-5	3	0,6	-22,91	-2,47	0,01899	0,866
110	-4	23	4,6	-17,91	-1,93	0,06213	2,832
115	-3	32	6,4	-12,91	-1,39	0,15189	6,923
120	-2	57	11,4	-7,91	-0,85	0,27789	12,667
125	-1	64	12,8	-2,91	-0,31	0,38008	17,325
130	0	111	22,2	2,09	0,23	0,38814	17,692
135	1	67	13,4	7,09	0,76	0,29871	13,616
140	2	47	9,4	12,09	1,30	0,17137	7,811
145	3	13	2,6	17,09	1,84	0,07362	3,356
150	4	4	0,8	22,09	2,38	0,02358	1,075

Ecart probable : $0,6745 \times 9^{\circ},28 = 6^{\circ},259$.

Gisement d'Angleur.
[*Anthraconauta belgica* Auctorum.]

1. Angle α .

$N = 197$

$A = 38^{\circ},37$

$\frac{N}{\sigma} = 24,472$

$n = 12$

$\sigma = 8^{\circ},05$

$h = 5$

$\frac{\sigma}{A} = 0,2098$

$\frac{N}{\sigma \sqrt{2\pi}} = 9,76$

$\alpha = x_i$	X_i	$\varphi(x_i) h$	$\varphi(x_i)$	$x_i - A$	$t = \frac{x_i - A}{\sigma}$	$\frac{1}{\sqrt{2\pi}} e^{-\frac{t^2}{2}}$	$\frac{N}{\sigma \sqrt{2\pi}} e^{-\frac{t^2}{2}}$
17	-4	1	0,2	-21,37	-2,65	0,01200	0,29
22	-3	3	0,6	-16,37	-2,00	0,05399	1,32
27	-2	17	3,4	-11,37	-1,4	0,14973	3,66
32	-1	41	8,2	-6,37	-0,7	0,31225	7,64
37	0	63	12,6	-1,37	-0,2	0,39104	9,57
42	1	38	7,6	+ 3,63	0,4	0,36827	9,01
47	2	15	3,0	+ 8,63	1,1	0,21785	5,33
52	3	9	1,8	+13,63	1,7	0,09405	2,30
57	4	7	1,4	+18,63	2,3	0,02833	0,69
62	5	1	0,2	+23,63	2,9	0,00595	0,15
67	6	—	—	+28,63	3,6	0,00061	0,01
72	7	2	0,4	+33,63	4,2	0,00006	—

Ecart probable : $0,6745 \times 8^{\circ},05 = 5^{\circ},43$.

2. Angle β .

$N = 193$

$A = 129^{\circ},062$

$\frac{N}{\sigma} = 18,2075$

$n = 12$

$\sigma = 10^{\circ},6$

$h = 5$

$\frac{\sigma}{A} = 0,0821$

$\frac{N}{\sigma \sqrt{2\pi}} = 7,26$

$\beta = x_i$	X_i	$\varphi(x_i) h$	$\varphi(x_i)$	$x_i - A$	$t = \frac{x_i - A}{\sigma}$	$\frac{1}{\sqrt{2\pi}} e^{-\frac{t^2}{2}}$	$\frac{N}{\sigma \sqrt{2\pi}} e^{-\frac{t^2}{2}}$
103	-5	5	1	-26,06	-2,5	0,01753	0,32
108	-5	8	1,6	-21,06	-2,0	0,05399	0,98
113	-3	5	1	-16,06	-1,5	0,12952	2,35
118	-2	20	4	-11,06	-1,0	0,24197	4,41
123	-1	19	3,8	-6,06	-0,57	0,33877	6,17
128	0	51	10,2	-1,06	-0,1	0,39695	7,23
133	1	38	7,6	3,94	0,4	0,36827	6,71
138	2	25	5	8,94	0,8	0,28969	5,27
143	3	9	1,8	13,94	1,3	0,17137	3,12
148	4	9	1,8	18,94	1,8	0,07895	1,44
153	5	3	0,6	23,94	2,3	0,02833	0,52
158	6	1	0,2	28,94	2,7	0,01042	0,19

Ecart probable : $0,6745 \times 10^{\circ},6 = 7^{\circ},15$.

Gisement de Stud.

[*Anthraconauta belgica* Auctorum.]1. Angle α .

$N = 149$

$A = 37^{\circ},53$

$\frac{N}{\sigma} = 19,426$

$n = 14$

$\sigma = 6^{\circ},67$

$h = 3$

$\frac{\sigma}{A} = 0,178$

$\frac{N}{\sigma \sqrt{2\pi}} = 7,750$

$\alpha = x_i$	X_i	$\varphi(x_i)h$	$\varphi(x_i)$	$x_i - A$	$t = \frac{x_i - A}{\sigma}$	$\frac{1}{\sqrt{2\pi}} e^{-\frac{t^2}{2}}$	$\frac{N}{\sigma \sqrt{2\pi}} e^{-\frac{t^2}{2}}$
21	-5	1	0,33	-16,53	-2,48	0,01850	0,359
24	-4	5	1,67	-13,53	-2,03	0,05099	0,990
27	-3	5	1,67	-10,53	-1,58	0,11464	2,227
30	-2	18	6	-7,53	-1,13	0,21075	4,094
33	-1	8	2,67	-4,53	-0,68	0,31644	6,147
36	0	42	14	-1,53	-0,23	0,38814	7,540
39	1	35	11,67	1,47	+0,22	0,38911	7,559
42	2	8	2,67	4,47	0,67	0,31854	6,188
45	3	15	5	7,47	1,12	0,21312	4,140
48	4	2	0,67	10,47	1,57	0,11650	2,263
51	5	8	2,67	13,47	2,02	0,05199	1,010
54	6	—	—	16,47	2,47	0,01899	0,369
57	7	—	—	19,47	2,92	0,00565	0,110
60	8	2	0,67	22,47	3,37	0,00138	0,027

Ecart probable : $0,6745 \times 6,67 = 4,50$.2. Angle β .

$N = 147$

$A = 135^{\circ},99$

$\frac{N}{\sigma} = 14,86$

$n = 10$

$\sigma = 9^{\circ},89$

$h = 5$

$\frac{\sigma}{A} = 0,073$

$\frac{N}{\sigma \sqrt{2\pi}} = 5,9296$

$\beta = x_i$	X_i	$\varphi(x_i)h$	$\varphi(x_i)$	$x_i - A$	$t = \frac{x_i - A}{\sigma}$	$\frac{1}{\sqrt{2\pi}} e^{-\frac{t^2}{2}}$	$\frac{N}{\sigma \sqrt{2\pi}} e^{-\frac{t^2}{2}}$
110	-5	1	0,2	-25,99	-2,63	0,01264	0,188
115	-4	5	1,0	-20,99	-2,12	0,04228	0,268
120	-3	7	1,4	-15,99	-1,62	0,10755	1,598
125	-2	15	3,0	-10,99	-1,11	0,21548	3,202
130	-1	24	4,8	-5,99	-0,61	0,33112	4,920
135	0	28	5,6	-0,99	-0,10	0,39695	5,899
140	1	29	5,8	4,01	0,41	0,36665	5,448
145	2	18	3,6	9,01	0,91	0,26367	3,918
150	3	16	3,2	14,01	1,42	0,14569	2,165
155	4	4	0,8	19,01	1,92	0,06332	0,941

Ecart probable : $0,6745 \times 9,89 = 6,67$.

F. — CONCLUSIONS BIOLOGIQUES RELATIVES A LA « VARIABILITE INDIVIDUELLE » ET AUX « VARIATIONS SPECIFIQUES ».

Les conclusions biologiques de l'étude biométrique des populations d'*Anthraconauta* des gisements A. — Angleur, A. — Stud et A. — Wérister doivent montrer pour chacune d'elles l'existence ou l'absence, d'une part, de corrélation entre les caractères mesurés et la « variabilité individuelle » de ceux-ci, d'autre part, de « formes variantes », dites spécifiques. A l'instar de P. PELSENEER ⁽⁶⁴⁾, nous considérons la « variabilité morphologique individuelle » comme résultant d'un ensemble de caractères qui, étant communs à tous les individus, sont donc normaux. Ces formes sont distinctes des « variétés spécifiques » portant, elles, des modifications d'ordre plus élevé. Rappelons encore que cette variabilité est, comme toujours en paléontologie, celle de la coquille, organe extérieur de l'animal, et qu'elle ne reflète pas la nature des variations correspondantes de l'organisme. Variabilité individuelle et, éventuellement, variations spécifiques doivent être examinées en fonction de l'âge (ou de la taille) et du point de vue de l'intensité, de l'orientation et de la fréquence.

La croissance ou « variation normale » de l'organisme détermine des changements de la forme de la coquille, grands et intenses dans les premières phases du développement et qui diminuent de l'état embryonnaire à l'état adulte. Il convient donc d'essayer de situer la place qu'occupent les *Anthraconauta* dans la suite ontogénique des formes du groupe auquel elles appartiennent, aux fins de retrouver éventuellement leur type morphologique à l'un des stades de cette suite. Par là on pourrait retrouver ainsi dans les phases de cette série le sens des variations possibles.

Le développement ontogénique général des Mytilidés, auxquels nous rapportons les *Anthraconauta*, a été relativement bien étudié. Le rapprochement du genre avec des modiolomorphes ou des mytilimorphes n'a pas, à priori, une signification phyllétique par la similitude morphologique, mais il représente au moins un caractère de convergence très net. Ce phénomène, bien connu, est caractéristique, pour de nombreux groupes, du passage de la structure isomyaire à la structure anysomyaire chez les formes byssogènes, fixées en position pleurothétique [R. ANTHONY ⁽⁶⁵⁾]. L'embryogénie des espèces du groupe des Mytilidés a été étudiée par R. J. JACKSON ⁽⁶⁶⁾, F. BERNARD ⁽⁶⁷⁾, R. ANTHONY ⁽⁶⁸⁾, T. C. NEL-

⁽⁶⁴⁾ PELSENEER, P., 1919.

⁽⁶⁵⁾ ANTHONY, R., 1905, pp. 230 et suiv.

⁽⁶⁶⁾ JACKSON, R. T., 1890, pp. 363-364.

⁽⁶⁷⁾ BERNARD, F., 1898, pp. 67-86.

⁽⁶⁸⁾ ANTHONY, R., 1905, pp. 244 et suiv.

SON ⁽⁶⁹⁾ et d'autres encore. Elle a montré, dans le développement de la coquille embryonnaire, les phases suivantes, schématisées par la figure 11 :

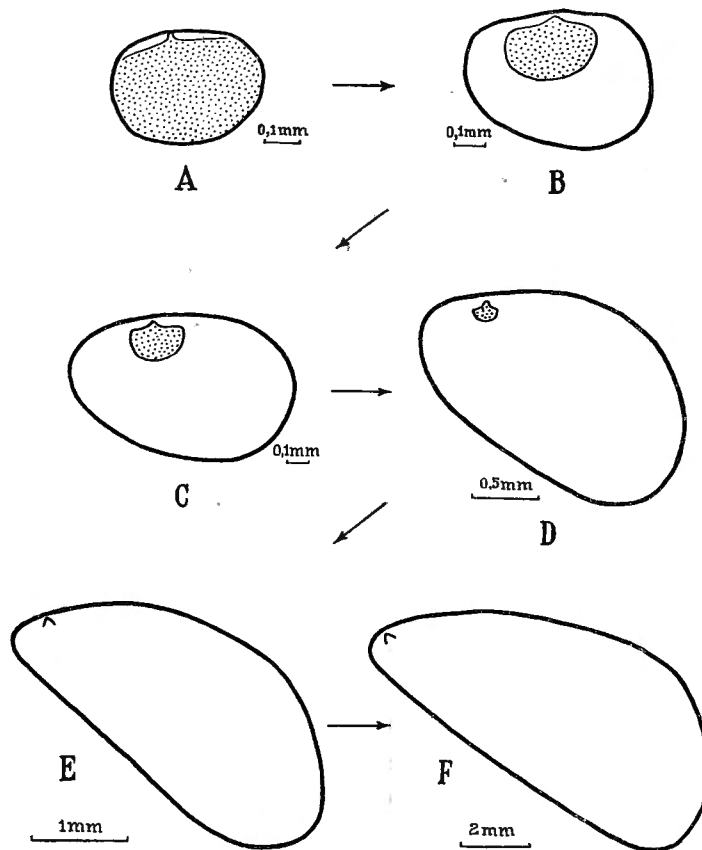


FIG. 11. — *Mytilus edulis* LINNÉ.

A. Prodysoconque définitive. — B. Dysoconque népionique.
C. Coquille au stade néologique. — D. Coquille au stade « modioloïde ».
E et F. Stades mytilloïdes.

Dès le début de la croissance de la dysoconque ou coquille vraie, au stade népionique, celle-ci développe une asymétrie progressive, passant par une phase de modiolisation, pour prendre ensuite le type mytilimorphe.

Ce processus apportant à la coquille son aspect définitif s'effectue par :

1° un allongement de la valve du côté postérieur, le côté antérieur s'accroissant moins vite, sans être cependant frappé d'avortement;

2° un accroissement du bord cardinal vers l'extrémité postérieure de la valve, ou, relativement, déplacement de l'umbo vers l'angle antérieur;

3° un déplacement angulaire de la position du diamètre de croissance maximum (d) par réduction de la valeur angulaire de l'angle α , ou, relativement, allongement oblique de la coquille.

⁽⁶⁹⁾ NELSON, T. C., 1928, pp. 180-190.

Ce processus de différenciation graduelle de la dyssocoque, aux stades népionique et néalogique, est propre à différents phyllums d'anysomyaires byssogènes. Rencontré dans des groupes divers, il ne signifie pour les uns qu'un cas de convergence; pour d'autres, il est le film d'une longue évolution phyllogénétique. Nous sommes enclins à considérer le type anthraconautiforme comme le représentant d'un stade de l'ontogénèse des Mytilidés, très voisin de l'état définitif auquel ont abouti les *Modiola* (voir chapitre I, § 2). Les variations « spécifiques », et à plus forte raison les variations « individuelles », du type anthraconautiforme apparaissent être déterminées principalement par la fluctuation des valeurs angulaires α et β et de la valeur de l'angle de mytilisation [R. ANTHONY ⁽⁷⁰⁾]; corrélativement, elles entraînent des modifications des valeurs indicielles.

Les conclusions de l'étude biométrique des trois populations A. — Angleur, A. — Stud, A. — Wérister sont également valables pour chacune d'elles. Les voici :

a) Il y a une bonne corrélation entre le caractère d (diamètre de croissance maximum, exprimant la taille, ou l'âge et les caractères a (hauteur), b (longueur), g (longueur du bord cardinal). Les valeurs élevées du coefficient de corrélation (voir tableau, p. 75) sont voisines de 1. Les valeurs les moins élevées, 0,864 et 0,886 pour ρ_{da} de A. — Angleur et ρ_{db} de A. — Stud sont à attribuer à une dispersion assez forte provenant de valeurs de a et g éloignées des moyennes. On peut donc conclure à l'existence de relations linéaires entre les caractères précités. Les rapports entre éléments de même nom de figures semblables étant constants, les indices caractéristiques $100 \frac{d}{a}$, $100 \frac{a}{b}$, $100 \frac{b}{g}$... sont indépendants du développement normal de la coquille. De la plus petite taille observée (2 mm.) à la plus grande (18,3 mm.), la coquille, au cours du développement, reste semblable à elle-même; sa croissance est homothétique dès l'adolescence. H. BOLTON ⁽⁷¹⁾, W. HIND ⁽⁷²⁾, E. DIX et A. E. TRUEMAN ⁽⁷³⁾ avaient noté que les petites coquilles de *A. minima*, *A. laevis* et *A. phillipsi*, grandes de 2 à 4 mm., paraissent généralement à contour plus arrondi et à plus faible obliquité. L'impression de cet état, que nous avons eue nous-mêmes, semble, après l'examen biométrique de petites formes analogues, n'être qu'une illusion de la vision, caractéristique de l'observation de petits segments de droites et de courbes.

b) L'absence de corrélation entre les caractères angulaires α , β et le caractère linéaire d se conçoit aisément, puisque, d'une part (au a), on constate le développement homothétique des coquilles au cours de la croissance et que,

⁽⁷⁰⁾ ANTHONY, R., 1905, p. 234.

⁽⁷¹⁾ BOLTON, H., 1915, pp. 676-677.

⁽⁷²⁾ HIND, W., 1894, p. 272.

⁽⁷³⁾ DIX, E. et TRUEMAN, A. E., 1931, p. 188.

d'autre part, les distributions de fréquences des variables α et β montrent une concentration nette des mesures obtenues autour de certaines valeurs.

c) L'examen du tableau de corrélation et des graphiques de distribution de fréquence des indices dénote, dans les limites observées des tailles d'individus, comme on pouvait s'y attendre, qu'il n'y a pas de discontinuité de croissance. Ceci, compte tenu des stases ou des ralentissements de la croissance dus aux phénomènes saisonniers et aux périodes d'activités reproductrices

d) L'examen du tableau de distribution de fréquence des indices caractéristiques et des angles α et β met en lumière que la « variabilité individuelle », toujours existante, présente une dispersion normale, régulière. Les courbes de fréquences n'offrant pas d'asymétrie, il n'y a pas de sens observable à cette variabilité, pas plus qu'on n'y peut reconnaître une amplitude ou une multiplicité. En un mot, l'allure de la « variabilité individuelle » est normale et il n'y a de « variétés spécifiques » dans aucun des trois gisements étudiés.

e) La dispersion normale des indices prouve que ces individus fossiles examinés ont été récoltés dans leur habitat naturel, à l'état grégaire, et n'ont pas été rassemblés en gisement, tels certains « mussels band », par action mécanique du milieu.

**G. — COMPARAISON DES INDICES TYPES DES DIFFÉRENTES
POPULATIONS POUR LA RECHERCHE D'UNE CONSPÉCIFICITÉ
ÉVENTUELLE DE LEURS ÉLÉMENTS.**

La comparaison de la spécificité de populations limniques ou dulcicoles, appartenant à des gisements différents, appelle le souvenir de cette remarque très générale de P. PELSENEER : « La variabilité est inégale entre des individus ou groupes d'individus dans une même espèce; elle est la plus grande dans les régions où les conditions d'existence sont les plus variées ». Elle est encore d'autant plus marquée que l'aire de dispersion de l'espèce est plus grande. C'est bien là le cas des *Anthraconauta* du terrain houiller étudiées. Cependant les individus des colonies étudiées ici et provenant de trois gisements distincts et éloignés présentent une conspécificité très nette.

L'ajustement des séries statistiques, obtenues suivant la loi de répartition de GAUSS, s'étant montré très satisfaisant, on a fait usage d'une propriété des écarts quadratiques, propre à la loi précitée, et permettant de définir l'intervalle quasi certain d'apparition des valeurs. On a admis aussi que la moyenne est la valeur la plus probable et que la « variable moyenne » définie plus haut (p. 27) obéit à la loi. Cette façon de faire est peut-être plus légitime que d'ajuster la série observée suivant la variable statistique elle-même par une répartition de GAUSS, qui n'est vraie que d'une façon asymptotique.

La « variable moyenne » correspondant à un indice caractéristique d'une espèce varie de façon quasi certaine dans un intervalle dont on peut déterminer l'ampleur. Cet intervalle est d'ailleurs d'autant plus petit que le nombre m de mesures est grand; il varie comme $\frac{1}{\sqrt{m}}$. Si des mesures analogues sont effectuées sur plusieurs indices caractéristiques d'une même espèce, on détermine chaque

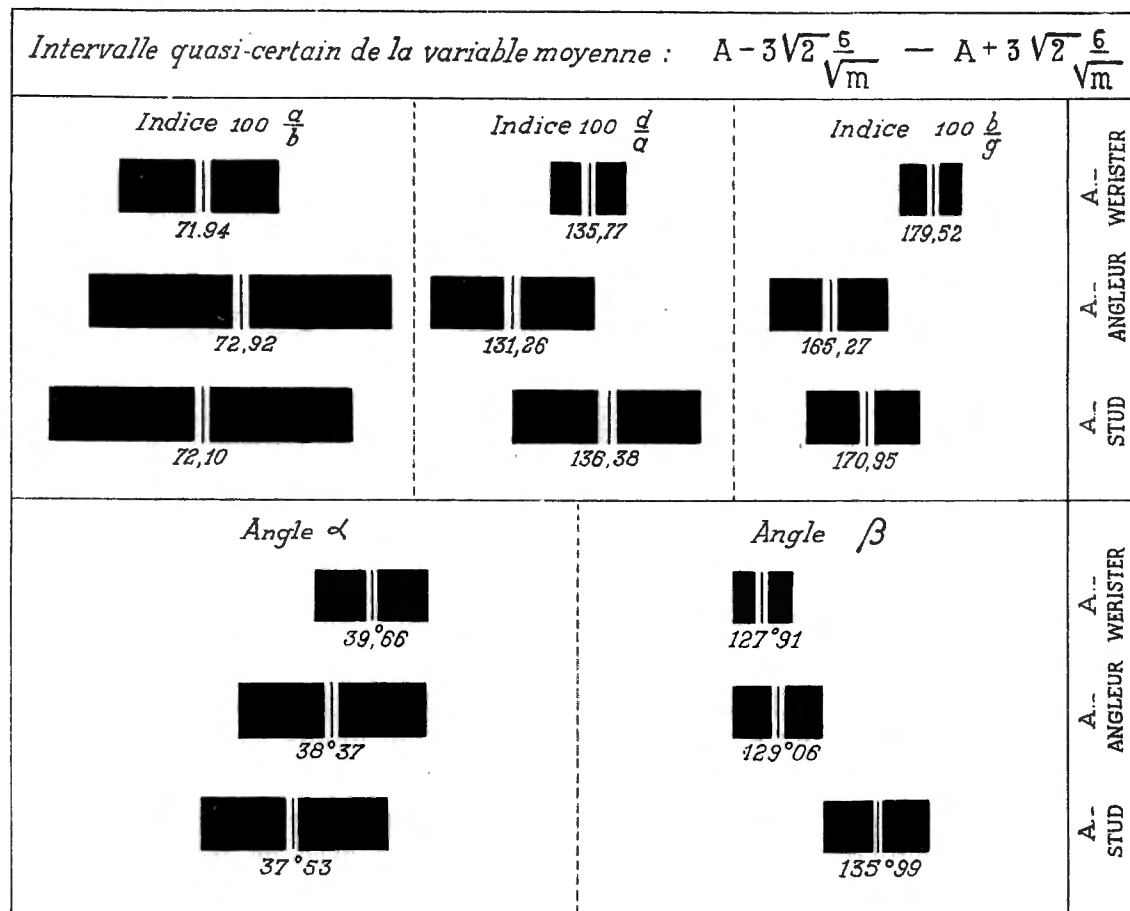


FIG. 12.

fois l'intervalle certain de leur apparition. Dès lors, si des mesures, portant sur d'autres espèces, révèlent des valeurs moyennes situées dans ces intervalles, il est permis de conclure, avec une probabilité d'autant plus grande que m est plus grand, à l'identité de ces espèces.

Les résultats obtenus sur les trois populations étudiées d'*A. belgica* (Auctorum) des gisements d'Angleur et de Stud et d'*A. minima*, du gisement de Wérister, espèce de référence, ont été groupés en tableaux relatifs, chaque fois à un indice ou à un angle caractéristiques.

De façon générale, A et σ étant la moyenne et la dispersion de l'indice ou de l'angle étudié, l'intervalle quasi-certain de la variable moyenne est

$$A - 3 \sqrt{2} \frac{\sigma}{\sqrt{m}}, \quad A + 3 \sqrt{2} \frac{\sigma}{\sqrt{m}}.$$

Indice	Espèce ou origine	Valeur moyenne	Dispersion	m	Intervalle quasi-certain
$100 \times \frac{a}{b}$	A. — Wérister	71,94	10,76	433	de 69,75 à 74,13
	A. — Angleur	72,92	14,10	209	de 68,78 à 77,06
	A. — Stud	72,10	12,54	150	de 67,75 à 76,45
$100 \times \frac{d}{a}$	A. — Wérister	135,77	13,38	433	de 133,05 à 138,49
	A. — Angleur	131,26	21,04	209	de 125,08 à 137,44
	A. — Stud	136,98	17,28	150	de 130,99 à 142,97
$100 \times \frac{h}{g}$	A. — Wérister	179,52	20,51	425	de 175,30 à 183,74
	A. — Angleur	165,27	26,00	201	de 157,47 à 173,07
	A. — Stud	170,95	22,24	146	de 163,14 à 178,76
Angle	Espèce ou origine	Valeur moyenne	Dispersion	m	Intervalle quasi-certain
α	A. — Wérister	39°,66	6°,26	421	de 38°,37 à 40°,95
	A. — Angleur	38°,37	8°,05	197	de 35°,94 à 40°,80
	A. — Stud	37°,53	6°,67	149	de 35°,20 à 39°,86
β	A. — Wérister	127°,91	9°,28	423	de 125°,99 à 129°,83
	A. — Angleur	129°,06	10°,60	193	de 125°,85 à 132°,27
	A. — Stud	135°,99	9°,89	147	de 132°,52 à 139°,46

L'examen comparatif des valeurs moyennes du tableau, transcrites sur le graphique de la figure 12, est déjà significatif : la distance maximum entre deux de ces valeurs est de l'ordre de grandeur maximum de 8 %; les intervalles quasi-certains des « valeurs moyennes » se chevauchent au moins largement.

Il est remarquable d'arriver à une concordance telle avec des populations peu nombreuses, et, bien que l'argument des grands nombres ne puisse être invoqué de manière incontestable, on peut conclure avec de très fortes présomptions à l'identité entre les espèces étudiées, à savoir entre l'*Anthraconauta belgica* HIND et l'*Anthraconauta minima* (HIND, non LUDWIG).

CONCLUSIONS GÉNÉRALES.

Si l'étude biométrique des organismes fossiles ne révèle rien sur la nature ou le principe des phénomènes biologiques, « elle permet cependant de résoudre des problèmes de spécificité difficilement solubles », et aussi de préciser les limites de la « variation spécifique ».

« Pour que les méthodes statistiques donnent les résultats que nous sommes en droit d'en attendre, il est nécessaire qu'un nombre considérable de documents

soient accumulés; il est donc désirable que les études sur ce sujet se multiplient. Mais, pour que les résultats de ces études soient comparables, il est indispensable qu'à l'instar des anthropologistes, les paléontologues standardisent leurs méthodes » (M. GLIBERT, 1945, p. 4). Au préalable, celles-ci doivent être exposées, confrontées et expérimentées, ainsi que précisées dans leur signification mathématique profonde et dans le sens réel de leurs possibilités pratiques, à la fois par le biologiste et le mathématicien, pour justifier et ordonner pleinement leur emploi.

Au surplus, dans de nombreux cas, les grandeurs de l'objet lui-même, qui deviennent mesures et sont considérées comme caractères morphologiques, demandent à être examinées auparavant dans leur valeur biologique et dans leur valeur géométrique. Sur ce dernier point, force nous est de constater que les caractères d'une coquille (tels que la longueur, la hauteur, etc.), que les zoologistes estiment représentatifs, sont parfois dépourvus du sens précis que leur conférerait une définition ou une propriété géométrique rigoureuse. Mesurer des « longueurs » ou des « largeurs » est une entreprise raisonnable, à condition qu'on sache définir exactement ces grandeurs. On ne mesure pas de façon bien précise une « longueur » de coquille, quand on ne sait pas exactement où elle commence ni où elle finit, étant donné que le segment de droite qui la représente est limité par des arcs. Or, c'est cela qu'on se contente de faire le plus souvent.

A notre avis, à des caractères difficiles à mesurer convenablement, on aurait tout intérêt à substituer des caractères définis sans ambiguïté par des repères géométriques simples. En nombre suffisant, ceux-ci peuvent permettre, par l'application de méthodes statistiques, de déterminer d'abord l'équation de la courbe du bord libre de la coquille, puis l'équation de sa surface géométrique. Ce dernier problème a déjà été traité par M. LISON (1942, pp. 377 et suiv.). Dans un travail ultérieur nous reviendrons sur l'application de ces méthodes qui apparaissent rationnelles.

Quoi qu'il en soit, même avec les imperfections résultant et de l'état de conservation du matériel étudié, et des définitions de certains caractères mesurés, l'étude biométrique de l'*Anthraconauta belgica* HIND et de l'*Anthraconauta minima* HIND (non LUDWIG) a démontré une fois de plus qu'on peut arriver à des conclusions à probabilités très élevées qui, pour le biologiste, ont la valeur de certitudes.

INDEX BIBLIOGRAPHIQUE.

- AMALIZKY, W., 1892, *Ueber die Anthracosien der Permformation Russlands*. (Palaeontographica, Stuttgart, Bd. 39, pp. 125-212.)
- ANCIEN, CH. et VAN LECKWIJCK, W., 1942, *Découverte d'une faune namurienne dans la vallée de la Berwinne*. (Bull. Acad. Roy. Belg., Bruxelles, Cl. Sci., 5^e sér., t. XXVIII, pp. 889-893.)
- ANTHONY R., 1905, *Influence de la fixation pleurothétique sur la morphologie des mollusques acéphales dimyaires*. (Ann. Sci. Nat., Paris, 9^e sér., t. I, zool., pp. 165-338, 57 fig.)
- BERNARD, F., 1895, *Éléments de Paléontologie*. (Paris, Baillière, 1.100 p.)
- 1895-1897, *Notes sur le développement et la morphologie de la coquille chez les Lamellibranches*. (Bull. Soc. Géol. France, Paris, première note, vol. XXIII, pp. 104-155, 1895; deuxième note [Taxodontes], vol. XXIV, pp. 54-82, 1896; troisième note [Anysomyaires], vol. XXIV, pp. 412-449, 1896; quatrième et dernière note, vol. XXV, pp. 559-566, 1897.)
- 1898, *Recherches ontogéniques et morphologiques sur la coquille des Lamellibranches*. (Ann. Sci. Nat., Paris, 8^e sér., t. VIII, zool., pp. 1-201, 12 pl., 17 fig.)
- BOLTON, H., 1915, *The fauna and stratigraphy of the Kent Coalfield*. (Trans. Manchester Geol. Soc., vol. XXXIV, 60 p., 3 pl.)
- DALL, H. in ZITTEL, K. and EASTMAN, CH., 1927, *Textbook of Paleontology*. (Mac Millan & C^o, London, vol. I, 822 p., 1594 fig.)
- DARMOIS, G., 1941, *Statistique et applications* (2^e édition). (A. Colin, Paris, 200 p.)
- DAVIES, J. H. and TRUEMAN, A. E., 1927, *A revision of the non marine Lamellibranchs of the Coal measures*. (Quart. Journ. Geol. Soc. London, vol. LXXXIII, pp. 210-259, pl. 16-17.)
- DELTHEIL, R., 1930, *Erreurs et moindres carrés*. (Gauthier-Villars, Paris, 161 p.)
- DEMANET, F., 1941, *Faune et stratigraphie de l'étage namurien de la Belgique*. (Mém. Mus. roy. Hist. nat. Belg., Bruxelles, n^o 97, 298 p., 18 pl.)
- 1943, *Les Horizons marins du Westphalien de la Belgique et leurs faunes*. (Mém. Mus. roy. Hist. nat. Belg., Bruxelles, n^o 101, 161 p., 9 pl.)
- DEWAR, W., 1939, *Anthraconauta (Anthracomya) minima (auctorium) and its associates in the Lancashire coalfield*. (Bull. Geol. Survey of Gr. Brit., London, n^o 1, pp. 46-65, 2 fig., pl. IV.)
- DIX, E. and TRUEMAN, A. E., 1931, *Some non-marine Lamellibranchs from the upper part of the coal-measures*. (Quart. Journ. Geol. Soc., t. LXXXVII, pp. 180-211.)
- DOUVILLE, H., 1912, *Classification des Lamellibranches*. (Bull. Soc. Géol. France, Paris, 4^e sér., t. XII, pp. 419-467, 1 table, 69 fig.)
- 1925, *Comment ont apparu certaines formes nouvelles, Rudistes et Chames, Mytilus et Dreissensia, Anomia et Paranomia*. (C. R. Acad. Sci. Paris, t. CLXXII, pp. 887-892.)

- FEDOTOV, D. M., 1937, *The Lamellibranchiata from the Coal-Bearing deposits of the Kuznetsk Basin and their stratigraphical significance*. (Trans. Central Geol. and Prospect. Inst., fasc. 97; pp. 5-96.)
- 1938, *The pelecypoda from the Coal-Bearing deposits chiefly from the industrial district of the Kuznetsk Basin*. (Bull. Acad. Sci. U.R.S.S., Moscou, Cl. Sci. math. et nat., sér. biol., n° 1, pp. 219-260.)
- FELDMAN, W. M., 1935, *Biomathematics*. (Ch. Griffin, London, 2^e édit., 480 p., 169 fig.)
- FIRKET, A., 1875, *Modiola du schiste houiller d'Angleur*. (Ann. Soc. Géol. Belg., Liège, t. II, pp. 162-163.)
- HAAS, F., 1941, in BRONNS, H. C., *Klassen und ordnung des Tierreichs. Bivalvia*. (Leipzig, teil II, lief. 3.)
- HENDERSON, J., 1935, *Fossil non-marine of North-America*. (Geol. Soc. of America, New-York, Spec. Paper, n° 3, 291 p.)
- HIND, W., 1893, *On the affinities of Anthracoptera and Anthracomya*. (Quart. Journ. Geol. Soc. London, vol. LIX, pp. 249-275, pl. VII-X.)
- 1894-1896, *A monograph on Carbonicola, Anthracomya and Naiadites*. (Palaeontographical Soc., London, vol. I, pp. 1-80, 11 pl.; vol. II, pp. 81-170, pl. XII-XX; vol. III, pp. 171-182, pl. XXI.)
- 1911, *Les Faunes conchyliologiques du terrain houiller de Belgique*. (Mém. Mus. roy. Hist. nat. Belg., Bruxelles, t. VI, 18 p., 1 pl.)
- JACKSON, R. T., 1890, *Phylogeny of the Pelecypoda, the Aviculidea and their allies*. (Mem. Boston Soc. Nat. Hist., vol. IV, pp. 277-400.)
- LINDER, A., 1945, *Statistische methoden für Naturwissenschaftler, Mediziner und Ingenieure*. (Lehrbücher und Monographien aus dem Gebiete der exakten Wissenschaften, Basel, Verlag Birkhäuser, 150 p.)
- LISON, L., 1939, *Étude biométrique sur les Lamellibranches, sur la mécanique du développement des surfaces spirales logarithmiques des êtres vivants*. (Bull. Acad. roy. Belg., Bruxelles, 5^e sér., t. XXV, pp. 260-278.)
- 1940, *Analyse quantitative des facteurs de la morphogénèse des formes logarithmiques biologiques*. (Bull. Acad. roy. Belg., Bruxelles, 5^e sér., t. XXVI, pp. 64-82.)
- 1941, *Sur la mécanique du développement des formes turbinées*. (Bull. Acad. roy. Belg., Bruxelles, 5^e sér., t. XXVII, pp. 241-259.)
- 1942, *Caractéristiques géométriques naturelles des coquilles de Lamellibranches*. (Bull. Acad. roy. Belg., Bruxelles, 5^e sér., t. XXVIII, pp. 377-390.)
- MORLEY-DAVIES, A., 1933, *The Bases of classifications of the Lamellibranchiata*. (Proc. Malac. Soc., London, vol. XX, part 6, pp. 322-326.)
- NELSON, T. C., 1928, *Pelagic dissoconchs of the common mussel, Mytilus edulis, with observations on the behavior of the larvae of allied genera*. (Bull. Woods Hole Marine Biol. Labor., Lancaster, vol. LV, pp. 180-192.)
- NERINCK, E., 1943, *Applications de la méthode biométrique et du baguage dans l'étude écologique des Chéiroptères*. (Bull. Mus. roy. Hist. nat. Belg., Bruxelles, t. XXIX, n° 62, 16 p., 4 fig.)
- NEUMAYR, M., 1883, *Zur morphologie der bivalven schlosses*. (Sitzber. der K. Akad. des Wiss., Wien, Bd 88, 34 s., X pl.)

- NEUMAYR, M., 1891, *Beitrage zur einer morphologischen eintheilung der bivalven.* (Denkschr. K. K. Akad. Wiss., Wien, Math.-Nat. Klass., Bd. 58, s. 705-801.)
- NEWELL, N. D., 1942, *Late Paleozoic Pelecypods: Mytilacea.* (Univ. Kansas Public., Lawrence, vol. X, part 2, 78 p., 15 pl., 22 fig.)
- PAUL, H., 1941, *Lamellibranchiata infracarbonica.* (Foss. Cat., Animalia, Paris, 91, 438 p.)
- PEARSON, K., 1896, *Regression, Heredity and Panmixia.* (Philos. Trans. Roy. Soc., série A, vol. CLXXXVII, p. 253.)
- PELSENEER, P., 1893, *La formation de variétés chez la moule comestible.* (Bull. Soc. roy. malac. Belg., t. XXVIII, 4 p.)
- 1906, in RAY LANKASTER, *A Treatise on Zoology.* (London, Adam et Ch. Black, part 5, Mollusca, 347 p., 301 fig.)
- 1919, *Les variations et leur hérédité chez les Mollusques.* (Mém. Acad. roy. Belg., Bruxelles, Cl. Sci., t. V, 780 p., 279 fig.)
- 1935, *Essais d'éthologie zoologique d'après l'étude des Mollusques.* (Public. Fondat. Agathon de Potter, Bruxelles, Palais des Académies, n° 1, 642 p.)
- PRUVOST, P., 1913, *Les niveaux à Lamellibranches d'eau douce dans le terrain houiller du Nord de la France, leur faune et leur distribution stratigraphique.* (Ann. Soc. Géol. Nord, Lille, t. XLII, 46 p., 12 fig., 2 pl.)
- 1919, *Introduction à l'étude du terrain houiller du Nord et du Pas-de-Calais.* (Mém. pour servir à l'étude de la Carte géologique détaillée de la France, Paris, 584 p., 29 pl., 51 fig.)
- 1930, *La Faune continentale du terrain houiller de la Belgique.* (Mém. Mus. roy. Hist. nat. Belg., Bruxelles, t. XLIV, pp. 247-248.)
- RAGOZIN, L., 1933, *The geological distribution of Pelecypods in the productive series of the Kusnetsk Basin.* (Rec. of Geol. of the West Siber. Région Tomsk, 13, pp. 36-63.)
- 1935, *Die Pelecypoden der Balackonskfolge des Kusnetsk-beckens.* (Trudy Univ. Tomsk, 88, pp. 57-74, 4 pl.)
- RAY-LANKASTER, 1906, *A Treatise on Zoology.* (London, Adam et Ch. Black, part 5, Mollusca, 347 p., 301 fig.)
- REED, F. R. C., 1939, *Non-marine Lamellibranchs, .. from the Speckled Sandstone Formation (Punjabian) of the Salt Range.* (Rec. Geol. Survey India, vol. LXXIV, pp. 474-491.)
- RENIER, A., STOCKMANS, FR., DEMANET, F., VAN STRAELEN, V., 1938, *Flore et Faune houillères de la Belgique. Introduction à l'étude paléontologique du Terrain houiller.* (Patrim. Mus. roy. Hist. nat. Belg., Bruxelles, 312 p., 146 fig., 144 pl.)
- RISSER, R., 1932, *Applications de la statistique à la démographie et à la biologie.* (Traité du calcul des probabilités et de ses applications, par E. BOREL, Gauthier-Villars, Paris, t. III, 249 p.)
- RISSER, R. et TRAYNARD, C. E., 1933, *Les principes de la statistique mathématique.* (Gauthier-Villars, Paris, t. XI, 338 p.)
- SAINTE-LAGUE, A., 1932, *Probabilités et morphologie.* (Coll. Act. Scient. et Industr., Paris, n° 36, 30 p., 7 fig.)
- STAINIER, X., 1932, *Stratigraphie des assises inférieures du Bassin houiller du Hainaut.* (Jumet, Hosdain, 31 p., 151 pl.)

- STEINMAN, E. et DODERLEIN, L., 1890, *Éléments de Paléontologie*. (Leipzig, 840 p., 1.030 fig.)
- TRUEMAN, A. E. and WEIR, J., 1946-1947, *A monograph of British carboniferous non-marine lamellibranchia*. (Palaeontographical Soc., London, part I, II.)
- VAN DEUREN, P., 1934-1935, *Leçons sur le calcul des probabilités*, t. I. *La théorie des probabilités*, xvii+546 p., t. II. *Les applications des probabilités*, xvi+556 p. (Gauthier-Villars, Paris, et Wesmael Charlier, Namur.)
- WEHRLI, H., 1933, *Die carbonischen süßwasser fauna der zeche Baldur (Westfalen)*. (Neu Jahrb. Min Geol. Pal., Stuttgart, Beil Abt. B, 119, s. 172-188.)
- WEIR, J., 1944, *A review of recent work on the Permian non-marine Lamellibranchs and its bearing on the affinities of certain non-marine genera of the upper-Paleozoic*. (Trans. Geol. Soc. Glasgow, vol. XX, part 3, pp. 291-340.)
- WEIR, J. and LEITCH, D., 1936, *The zonal distribution of the non-marine Lamellibranchs in the coal-measure of Scotland*. (Trans. Roy. Soc. Edinburg, vol. LVIII, pp. 697-751, 4 pl., 15 fig.)
- WHITE KATHLEEN, M., 1932, *Mytilus*. (Mem. Liverpool Marine Biol. Comit., Liverpool, vol. XXI, 117 p., 10 pl.)
- YULE, U. and KENDALL, M. G., 1945, *An introduction to the theory of statistics*. (Ch. Griffin, London, 13^e édit., 570 p., 55 fig., 4 pl.)
- ZITTEL, K. V. and EASTMAN, CH. R., 1927, *Textbook of Paleontology*. (Mac Millan, London, vol. I, 822 p., 1.594 fig.)
-

TABLE DES MATIÈRES

	Pages
INTRODUCTION	3
CHAPITRE I.	
§ 1. Place des mollusques limniques ou dulcicoles du Terrain houiller dans la classification zoologique systématique	5
§ 2. Phyllogénèse des <i>Anthraconauta</i>	7
§ 3. Caractères généraux des <i>Anthraconauta</i>	8
§ 4. Spécificité d' <i>Anthraconauta belgica</i> HIND	9
§ 5. Nécessités d'une étude biométrique	10
§ 6. Position stratigraphique des gisements et provenance du matériel étudié ...	11
CHAPITRE II.	
INTRODUCTION	12
A. — Choix des caractères morphologiques à mesurer. Grandeur de la précision des mesures	13
B. — Choix des méthodes de la statistique mathématique à appliquer à l'étude biométrique	14
Séries statistiques à une variable.	
§ 1. Distribution de fréquences. Représentation géométrique	14
§ 2. Problème fondamental de l'ajustement des séries statistiques	17
§ 3. Moyennes et écarts. Moments	19
§ 4. Principe des moments	22
§ 5. La loi de GAUSS	23
§ 6. Propriétés de la loi de GAUSS	25
§ 7. De la moyenne arithmétique considérée comme variable statistique	27
§ 8. Symétrie et asymétrie d'une distribution observée	28
§ 9. Méthode pratique de calcul des moyennes et de la dispersion... ..	30
§ 10. Liaison entre deux variables statistiques. Corrélation	31
§ 11. Distribution de fréquences d'un complexe à deux variables statistiques ...	33
§ 12. Dépendance et indépendance au sens des probabilités	36

	Pages
§ 13. Moyennes liées. Lignes de régression	37
§ 14. Écart quadratique et rectangulaire. Coefficient de corrélation	38
§ 15. Signification du coefficient de corrélation	40
§ 16. Une loi particulière d'ajustement de répartitions à deux variables : la loi de BRAVAIS... ..	40
§ 17. Droites de régression d'une distribution quelconque	41
§ 18. Erreur systématique et écart quadratique du coefficient de corrélation ...	42
C. — Mesuration des individus (<i>Anthraconauta</i>)	44
CHAPITRE III.	
D. — Étude de la corrélation des caractères (<i>Anthraconauta</i>)	44
INTRODUCTION	44
MÉTHODES	45
REPRÉSENTATION GRAPHIQUE	46
§ 1. Tableaux de corrélations	47
§ 2. Calcul des moyennes	56
§ 3. Calcul des dispersions	57
§ 4. Calcul des écarts rectangulaires et des coefficients de corrélation et de régression	63
E. — Étude des indices caractéristiques et de leur distribution de fréquence (<i>Anthra- conauta</i>)	76
§ 5. Distribution de fréquences	77
§ 6. Étude des angles caractéristiques	83
F. — Conclusions biologiques relatives à la « variabilité individuelle » et aux « variations spécifiques »	87
G. — Comparaison des indices types des différentes populations pour la recherche d'une conspécificité éventuelle de leurs éléments... ..	90
CONCLUSIONS GÉNÉRALES	92
INDEX BIBLIOGRAPHIQUE	94
TABLE DES MATIÈRES... ..	98



