JOURNAL DE LA SOCIÉTÉ STATISTIQUE DE PARIS

J. DUFRÉNOY

Courbes de croissance

Journal de la société statistique de Paris, tome 113 (1972), p. 281-302
<http://www.numdam.org/item?id=JSFS_1972__113__281_0>

© Société de statistique de Paris, 1972, tous droits réservés.

L'accès aux archives de la revue « Journal de la société statistique de Paris » (http://publications-sfds.math.cnrs.fr/index.php/J-SFdS) implique l'accord avec les conditions générales d'utilisation (http://www.numdam.org/conditions). Toute utilisation commerciale ou impression systématique est constitutive d'une infraction pénale. Toute copie ou impression de ce fichier doit contenir la présente mention de copyright.

NUMDAM
Article numérisé dans le cadre du programme
Numérisation de documents anciens mathématiques
http://www.numdam.org/

COURBES DE CROISSANCE

En portant en ordonnées sur échelle arithmétique, les tailles (dimensions linéaires) ou les poids (ou les volumes) d'un individu aux divers temps (t) du cycle de croissance, les valeurs de (t) étant portées en abscisses, on définit des points (t, y) qui s'alignent sur une courbe en S dite sigmoïde; a) les temps (t) sont portés sur l'échelle arithmétique d'abscisses depuis t_0 pris comme origine jusqu'à t_a, marquant la fin de la période de croissance (acquisition de l'état adulte); la courbe débute par une branche ascendante de courbe exponentielle, tant que la vitesse (ou taux d'accroissement) augmente, pour atteindre une vitesse maximale lorsque au cours de la première moitié du cycle de croissance $\frac{(t_a - t_0)}{2}$ aura été atteinte la moitié de la taille maximale (y max) vers laquelle tend asymptotiquement la branche décroissante symétrique de la branche ascendante initiale.

La courbe sigmoïde symétrique par rapport à $\frac{(t_a - t_0)}{2}$ lorsque les (t) sont portés sur échelle arithmétique, correspond à l'intégrale de la courbe de Distribution normale.

Pour les phénomènes biologiques, c'est, plutôt que le modèle de distribution normale, celui de distribution log normale qui représente le modèle le plus vraisemblable; il suffit d'utiliser la variable auxiliaire $t' = \log t$ pour revenir à une distribution normale.

Dans l'équation de la courbe sigmoïde de croissance, t représente la variable indépendante et y la variable dépendante.

La relation fonctionnelle entre y et t, considérée comme autocatalytique ou logistique met en œuvre trois constantes K, α et β, telles que $y = \frac{K}{[1 + \exp. (\alpha - \beta t)]}$ ou $\frac{1}{y} = A + B \exp. (-\beta t)$ lorsque $A = \frac{1}{K}$ et $B = \exp. \left(\frac{\alpha}{K}\right)$, ce qui ne laisse qu'une seule constante non linéaire (β).

Appliquant cette équation à l'étude de la croissance de rats élevés à la Division of Animal Health CSIRO, Parkville VIC. P. J. Pahl (*Aust. J. Biol. Sci.*, 1969, 22, 1077-1080) a calculé pour les mâles, $1/y = 0{,}003407 + 0.1409 \exp. (-0{,}082\ t)$; pour les femelles, $1/y = 0.005800 + 0.1248 \exp (-0.082\ t)$

D'après cette fonction $y = f(t)$, les taux absolus de croissance au temps $t = t_1$ soit $\left[\frac{dy\ (t)}{dt}\right] t = t_1$ et à toute autre période $t = t_i$ sont indiqués dans le tableau I.

Analyse statistique du tableau I de Pahl

Les valeurs de (t) en jours, ou celles de (x) en semaines, étant également espacées, la relation fonctionnelle entre taux absolu de croissance (maintenant désigné par y) et x, peut être aisément calculée sous la forme d'équation polynomiale $\hat{y} = \bar{y} + BZ_1 + CZ_2 + DZ_3 + EZ_4 \ldots$ lorsque

$$\hat{y} = \frac{\Sigma (y)}{n}; \quad Z_1 = \lambda_1 (x - \bar{x})$$

$$Z_2 = \lambda_2 \left[(x - \bar{x})^2 - \frac{n^2 - 1}{12}\right]$$

$$Z_3 = \lambda_3 \left[(x - \bar{x})^2 - (x - \bar{x}) \frac{3\,n^2 - 7}{20}\right] \ldots$$

$$B = \frac{\Sigma (Z_1 y)}{\Sigma (Z_1)^2} \qquad C = \frac{\Sigma (Z_2 y)}{\Sigma (Z_2)^2} \qquad D = \frac{\Sigma (Z_3 y)}{\Sigma (Z_3)^2} \ldots$$

Calcul des équations polynomiales

Le tableau I d'après Pahl indique pour chaque période successive de 7 jours, c'est-à-dire pour les nombres x de semaines, de 1 à 10, les valeurs des taux absolus de croissance, désignées ici par y, pour les mâles et pour les femelles.

On a calculé, pour les valeurs $x = 1$ à $x = 10$, d'après les 10 valeurs de y, les équations polynomiales $\hat{y} = \bar{y} + BZ_1 + CZ_2 + DZ_3 + EZ_4 \ldots$ soit pour

les mâles $\hat{y} = 3.69 + 0.133\, Z_1 - 0.368\, Z_2 - 0.0608\, Z_3 + 0.016\, Z_4$

les femelles $\hat{y} = 2.21 - 0.0124 - 0.2445\, Z_2 - 0.032\, Z_3$

Comparativement pour les 8 valeurs de $x = 3$ à $x = 10$, on a calculé pour les mâle $\hat{y} = 4.288 - 0.00291\, Z_1 - 0.2764\, Z_2$.

Nous commentons ces équations ci-après.

Tableau I (d'après P. J. Pahl)

	Mâles									Femelles	
t	x	y	Z_1	Z_2	$\hat{y}$	x'	Z_1	Z_2	$\hat{y}$	y	$\hat{y}$
7	1	1.0	−9	6	0.29					1.0	0.86
14	2	1.6	−7	2	2.028					1.6	1.80
21	3	2.5	−5	−1	2.557	1	−7	7	2.372	2.3	2.51
28	4	3.8	−3	−3	4.485	2	−5	1	4.026	3.1	2.98
35	5	5.0	−1	−4	5.029	3	−3	−3	5.088	3.5	3.20
42	6	5.9	1	−4	5.295	4	−1	−5	5.674	3.4	3.18
49	7	5.9	3	−3	5.193	5	1	−5	5.660	2.8	2.90
56	8	5.0	5	−1	4.720	6	3	−3	5.114	2.1	2.39
63	9	3.7	7	2	4.253	7	5	1	3.997	1.4	1.63
70	10	2.5	9	6	2.679	8	7	7	2.332	1.0	0.63
		36.9								22.1	22.6

Mâles, $x = 1$ à $x = 10$

Pour les 10 valeurs $x = 1$ à $x = 10$, formant les termes d'une progression arithmétique, l'équation polynomiale

$$\hat{y} = \bar{y} + BZ_1 + CZ_2 + DZ_3 + EZ_3$$

avec $Z_1 = 2\,(x - \bar{x})$ où $\bar{x} = 5.5$

$$Z_2 = \frac{1}{2}\left[(x - \bar{x})^2 - \frac{100^2 - 1}{12}\right] = \frac{1}{2}\left[(x - 5.5) - \frac{99}{12}\right]$$

$$Z_3 = \frac{5}{3}\left[(x - \bar{x})^3 - (x - \bar{x})\,\frac{(3)\,(10)^2 - (-7)}{20}\right]$$

prend pour les mâles la forme

$$\hat{y} = 3.69 + 0.133\,Z_1 - 0.368\,Z_2 - 0.0608\,Z_3 + 0.016\,Z_4$$

On note la grande importance relative du coefficient de régression quadratique (en Z_2) pour lequel on calcule $F = 23$, correspondant à $p = 0.998$

Analyse de la variance (tableau I, Mâles) Pahl Mâles

Source de variation	*SC*	*DL*	*CM*	*F*
Totale $\Sigma\,(y - 3.69)^2$	29.2	9		
Due à régression				
Linéaire (0.133) (44.4)	5 905	1	5.905	
Quad. (0.368) (48.6)	17.885	1	17.885	$\frac{17.88}{0.77} = 23$
	23.790			
Erreur	5.4	7	0.77	

Mâles $x = 3$ à $x = 10$

Pour les 8 valeurs $x = 3$ à $x = 10$, les valeurs de y se distribuent à peu près symétriquement relativement à la valeur maximale 5.9 atteinte pour $x = 6$ et $x = 7$; on prend comme variable indépendante $x' = x - 2$, telle que les 7 valeurs de x' se distribuent par rapport à $x' = 6.5$; soit $Z_1 = (x' - 6.5)$

On calcule l'équation polynomiale

$$\bar{y}' = 4.288 - 0.00291\,Z_1 - 0.2766\,Z_2$$

remarquable en ce que la contribution de la régression linéaire (en Z_1) est négligeable, celle de la régression quadratique (en Z_2) rendant compte de la quasi-totalité de la variance totale, et correspondant à une valeur de F significative au niveau $p = 0.995$.

La fonction $y = f(x)$ peut donc se représenter en coordonnées arithmétiques par une parabole et en coordonnées logarithmiques par une droite.

Analyse de la variance, Pahl table 1

Source de variation	*SC*	*DL*	*CM*	*F*
Totale $\Sigma\,(y - 2\,888)^2$	14.38	7		
Due à regression				
Linéaire (— 0.00291) (— 0.5)	0.001455	1		
Quad. (— 0.2766) (— 46.5)	12.6619	1	12.66	$\frac{12\,66}{0.34} = 38$
Erreur	1.7	5	0.34	

Femelles

Pour les 10 valeurs $x = 1$ à $x = 10$, les taux absolus de croissance y se distribuent en coordonnées arithmétiques sur un arc de parabole dont l'équation

$$\hat{y} = 2.21 - 0.124\, Z_1 - 0.2445\, Z_2 - 0.032$$

comporte comme seul coefficient contribuant significativement à la variance générale, le coefficient quadratique (en Z_2) pour lequel on calcule $F = 157$, correspondant à une valeur de p supérieure à 0.9995.

Les points de coordonnées (log x, log y) s'alignent d'ailleurs sur une droite.

Analyse de la variance Pahl, tableau I, femelles

Source de variation	*SC*	*DL*	*CM*	*F*
Totale $\Sigma\ (y - 2.21)^2$	8.28	9		
Due à régression				
Linéaire (— 0.0124) (— 4.1)	0.05	1		
Quadr. (— 0.2445) (— 32.3)	7.88	1	7.88	$\frac{7.88}{0.05} = 157$
Erreur	0.35	7	0.05	

Taux relatifs de croissance

Aux taux absolus de croissance estimés à chaque temps $t = t_1$ par la formule $\left[\frac{dy\ (t)}{dt}\right]\ t = t_1$ (taux absolus que nous avons désignés par y dans le tableau I) Pahl compare les taux relatifs de croissance au temps $t = t_1$ en posant $y\ (t)^{-1} \left[\frac{dy\ (t)}{dt}\right]\ t = t_1$ lorsque $\left[\frac{dy\ (t)}{dt}\right]\ t = t_1$ représente la dérivée de $y\ (t)$ par rapport à t, évalué à $t = t_1$.

En introduisant ces formules dans les fonctions de croissance Pahl obtient des taux relatifs de croissance tels que le taux pour la femelle à l'âge t correspond à celui du mâle pour l'âge $t + 7$; femelle et mâle sont déphasés d'une semaine; le taux absolu de croissance atteignant le maximum vers le 38e jour pour la femelle, le 45e pour le mâle; ce maximum étant 3,5 g par jour pour la femelle, 5,9 pour le mâle.

Comparaisons des courbes de croissance pondérales pour deux lignées de rats. J. J. Curley & A. Pelas, *Laboratory Animal Care*, 19, n° 5, part II, pp. 716-719 oct. 1969).

Le tableau indique pour les produits de femelle d'origine Long-Evans (BLU : (*LE*) ou Sprague-Dawley (CAW : CFE (*SD*) les poids y g du nouveau-né à la naissance ($x = 0$ jour) puis aux jours $x = 4$, 8, 14 et 21 (sevrage).

Les 5 points de coordonnées ($x_1\ y$) s'alignent sur un arc de parabole pour *SD*; pour *LE* les points (0, 5.13, 4, 9.19), (8, 15.14) et (21, 43.97) s'alignent sur une parabole, passant au-dessous du point (14, 34.63).

L'arc de parabole doit pouvoir se transformer en droite sur papier log log.

La valeur $x = 0$ ne permet pas d'utiliser une échelle log des abscisses, mais faute de pouvoir utiliser la transformation $x' = \log x$, on peut utiliser la transformation $x' =$ arc tang x, qui avec un module convenable de l'échelle arc tangente, permet d'obtenir pour *SD* un alignement rectilinéaire des 5 points (arc tang x; y).

Pour *LE* la valeur de y correspondant à arc tang. 14 se situe au-dessus de la droite passant par les 4 autres points.

Par intrapolation on évalue pour $x = 16$ la valeur $y = 30$, pour *SD* et $\bar{y} = 36$ pour *LE*, ce qui permet de calculer l'équation polynomiale pour les valeurs de $x = 0$, 4, 8, 16, en utilisant les valeurs Z_1 et Z_2 du tableau I

TABLEAU I

Poids y g de rats depuis la naissance ($x = 0$ jour) jusqu'au sevrage ($x = 21$ jours) (d'après table I, Lab. Animal Care, 19, p. 717).

				Long. Evans				Sprague Dewley			
J	x	Z_1	Z_2	y	Z_1y	Z_2y	$\hat{y}$	y	Z_1y	Z_2y	$\hat{y}$
0	0	— 7	7	5.13	— 35.91	35.91	5.774	5.87	— 41.09	41.09	5.91
4	1	— 3	— 4	9.19	— 27.57	— 36.76	8.641	9.66	— 28.98	— 38.64	9.56
8	2	1	— 8	15.14	15.41	— 120.12	14.693	14.70	14.70	— 117.60	14.78
14				34.63				24.46			
(16)	4	9	5	36.00	324.00	190.00	36.353	30	270.00	150	29.88
21				43.97				40.62			
		140	154		275.60	69.03	65.46	60.23	214.63	34.85	60.23

On calcule les équations polynomiales :

$$LE\ \hat{y} = 16.365 + 1.968\ Z_1 + 0.455\ Z_2$$
$$SP\ \hat{y} = 15.057 + 1.533\ Z_1 + 0.2263\ Z_2$$

A la naissance ($x = 0$) le nouveau né *SD* a un poids $y = 5.87 \pm 0.05$, significativement plus élevé que celui du nouveau né *LE*, avec $y = 5.13 \pm 0.07$, mais au sevrage *LE*, le poids $y = 43{,}97 \pm 0.84$ est significativement plus élevé que pour *SD* avec $y = 40.62 \pm 0.70$

Analyse de la variance (Sprague Dewley)

Source de variation	*SC*	*DL*	*F*
Totale $\Sigma\ (y - 15.057)^2$	336.93	3	
Due à régression			
Linéaire (1.533) (214.63)	329.03	1	$\frac{7.88}{0.02} = 394$
Quad. (0.2263) (34.85)	7.88	1	
Erreur	0.02	1	

Pour 1 *DL* au numérateur et 1 *DL* au numérateur la valeur de F est 161 au niveau de probabilité $p = 0.95$ et 648 au niveau de probabilité $p = 0.975$.

Le modèle de regression quadratique peut donc être conservé comme vraisemblable.

Échelle arc tangente comme échelle des temps

I. La croissance étant évaluée d'après l'accroissement dy de taille linéaire, du poids de l'individu au cours de chaque durée dt, la courbe de croissance a l'allure d'une sigmoïde avec deux asymptotes en $y = 0$ et $y = Y$ max, taille atteinte à la fin du cycle de croissance, à $t = t_u$, avec point d'inflexion ou de symétrie de la courbe coordonnées ($t_u/2$; Y max/2).

L'anamorphose de la courbe sigmoïde en droite peut s'effectuer soit en cumulant les y; en transformant les y cumulés en pour cent de Y max; en portant chaque pourcentage

cumulé sur échelle de probabilité normale en ordonnées, à son rang d'abscisses t; soit en adoptant pour l'échelle des t l'échelle arc tangente.

Utilisation de l'échelle arc tangente comme échelle des temps, pour la comparaison de courbes de croissance.

Le même modèle mathématique ayant été adopté pour représenter la courbe de croissance d'un individu A et celle d'un individu B, la comparaison peut porter sur les valeurs des paramètres calculés pour A et pour B : il y a intérêt à réduire le nombre de ces paramètres, si possible à deux, en effectuant une anamorphose des courbes en droite; dans bien des cas cette anamorphose s'obtient aisément en adoptant, pour les temps (t) l'échelle arc. tang. qui admet une position finie pour $t = 0$ et les positions symétriques par rapport à $t = 0$ pour $+ t_i$ et $- t_i$: plusieurs cas peuvent se présenter :

1° La transformation arc tang avec choix du module approprié, transforme en droites les 2 courbes de croissance :

a) les deux droites se superposent;

b) elles sont parallèles ne différant que par le paramètre de position;

c) elles ont des pentes différentes.

2° Le choix d'un module qui permet de transformer l'une des courbes en droite, fait apparaître une cassure pour l'autre anamorphose.

$x' \left[\frac{2}{\pi}\right]$ arc tang. $\left[\frac{x}{(K\,\pi)}\right]$ où M et K sont des paramètres de valeurs positives; tout nombre réel x peut ainsi être transformé en un autre nombre réel x' tel que $- M < x' < M$; le module $(K\,\pi)$ est la valeur de x correspondant à $x' = 1/2\,M$; on ajuste le module en multipliant par un facteur constant C, convenablement choisi, les valeurs marquées sur une échelle arc tang standard telle que celle de l'Orbit Electric Company définie par

$$M = \left(\frac{2}{\pi}\right) M_{\max} \text{ arc tan } \frac{N}{120\,\pi}$$

L'anamorphose en droite s'obtient aisément par transformation arc tang de l'échelle de la variable indépendante (t) notamment pour la courbe exponentielle et la courbe d'intégration de la distribution normale (en utilisant pour la variable dépendante sur échelle arithmétique) ou de la distribution log normale (en utilisant une échelle log des ordonnées pour la variable dépendante).

Applications numériques

Comparaison des courbes de croissance pondérales des mâles et des femelles de rats (BLU : *LE*) du 2e au 24e mois.

Pour chaque période $x = 2$ à $x = 24$ le tableau indique (d'après tables 1 et 2 de *Lab. An. Care*, 19, pp. 747 et 748) les poids y_g; nous avons calculé les y cumulés, qui ont été transformés en % cumulés (n) relativement à un total de 10 % supérieur à y cumulé pour $x = 24$, soit 2 500 pour les mâles, 1 600 pour les femelles.

Du début ($x = 2$) à la fin ($x = 24$) de la période, un mâle a gagné 307 g, dont la moitié du 2e au 6e mois; une femelle a gagné 138 g dont la moitié du 2e au 6e mois.

Nous adoptons comme échelle du temps (x) en abscisses, l'échelle arc tangente avec un module tel que l'espace 2 à 6 mesuré en centimètres sur l'échelle des abscisses soit approximativement égal à la distance 6 à 24. En portant chaque valeur de n sur échelle de probabilité normale en ordonnées, à son rang d'abscisses sur l'échelle arc tangente on définit 5 points qui s'alignent sur une droite.

Lab. An Care 19 (5) Tables 1 & 2; Rats long Evans

	Mâles			Femelles		
x	y	y cum.	n	y	y cum.	n
2	261.65	261.65	10.5	210.06	210.06	13.1
4	388.41	650.06	26	269.53	479.59	30
8	524.71	1174.77	47	297.25	776.84	48.1
12	555.10	1729.95	69.2	324.31	1101.15	68.7
24	567.92	2297.87	90	348.00	1449.15	90.5
		2 500	100		1 600	100

Courbe de croissance extrapolée pour la période de gestation

L'échelle arc tangente permet de définir la position — 21 jours, symétrique de la position + 21 relativement au jour zéro de la naissance.

L'anamorphose de la courbe de croissance par utilisation d'échelle, arc tangente permet de définir une droite qui, prolongée vers le bas et vers la gauche, doit intercepter l'ordonnée élevée de — 21 jours à un niveau qui correspond à un poids voisin de zéro

Ex. I — Les pourcentages cumulés (n) en pourcent de 2 500 pour les mâles, Long Evans, portés en ordonnées sur échelle de probabilité normale, chacun à son rang d'abscisses sur échelle arc. tang., déterminent les points par lesquels on peut tracer une droite de régression interceptant l'ordonnée de rang — 21 au niveau 0,1 %.

II — Une extrapolation plus satisfaisante s'obtient par anamorphose des courbes de croissance de rats mâles et femelles du 7e au 70e jour d'après Pahl.

On a calculé pour chaque fin de semaine, le gain y, de poids corporel, d'après le gain de poids quotidien au cours de chaque période successive de 7 jours (de chaque semaine) d'après Pahl, tableau I, (*Austral. J. Biol. Sci.*, 1969, 22, 1079); les y cumulés ont été transformés en % de 350 pour les mâles, de 180 pour les femelles.

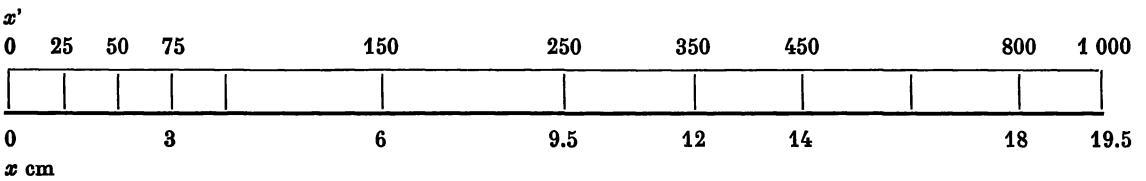
Mâles					Femelles		
J	S	y	y cum.	% de 350	y	y cum.	% de 180
7	1	7	7	2	7	7	4
14	2	11.2	18.2	5	11.2	18.2	10
21	3	17.5	35.7	11	16.1	34.3	18
28	4	26.6	62.3	19	21.7	56.0	31
35	5	35.0	97.3	28	24.5	80.5	41
42	6	41.3	138.6	40	23.8	104.3	58
49	7	41.3	179 9	51	19.6	123.9	68
56	8	35.0	214.9	61	14.7	138.6	77
63	9	25.9	240.8	68	9.8	148.4	82
70	10	17.5	258.3	75	6.3	154.7	85
			350	100		180	100

Les pourcentages cumulés (n) portés en ordonnées sur échelle de probabilité normale, chacun à son rang d'abscisses sur échelle arc tangente, définissant 10 points par lesquels passe une droite de régression, qui, prolongée vers la gauche et vers le bas, atteint un niveau compris entre 0.02 et 0.01 % à — 21 jours, début de la gestation.

III — Les anamorphoses par utilisation de l'échelle arc tangente pour les temps de $x = 5$ à $x = 90$ jours, pour les souris d'après Ogier et al. (*Rev. Path. Comp.*, 69 : Tableaux 1 et 2, p. 355, novembre-décembre 1969) donnent des droites comparables à celles obtenues pour les rats : les droites, extrapolées vers le bas et vers la gauche, interceptent, pour $x = -21$ jours, un niveau de l'échelle de probabilité normale inférieur à 0,02 % : pour les souris comme pour les rats, la courbe de croissance des femelles est déphasée d'environ une semaine relativement à celle des mâles.

L'Échelle arc tangente

L'échelle arc tangente peut être construite avec le double décimètre par exemple pour un module tel que les divisions x' sur l'échelle arc tangente soient distantes de l'origine $x = 0$, des valeurs x cm.



Références : Dufrenoy M. L. et Dufrenoy J. :

Utilisation de l'échelle arc tangente comme échelle des temps, *Jour. Soc. statist.*, Paris, 1954, p. 137.

L'échelle arc tangente et la représentation du temps biologique, *Ibid*, 1954, p. 144.

Anamorphose d'une courbe sigmoïde par utilisation de l'échelle arc tangente, *Ibid*, 1954, p. 146.

Le temps biologique, *Rev. Path. Gén.* et *Comp.*, n° 644, janvier 1955, pp. 129-138.

Taille optimale d'un être vivant

(V. A. Mezhzherin et A. A. Dyvldin, *Nature*, 227, 306-307, 18 juillet 1970).

La bioénergétique étudie les relations entre taux des échanges d'énergie et autres paramètres de l'organisme : Quant à la chaleur :

1) La production totale $Q(m)$ d'un organisme augmente avec la masse du corps nu selon la fonction $Q(m) \approx am^b$ avec b au plus égal à 1.

2) La production spécifique $q(m)$ décroît lorsque la masse m augmente, soit $q(m) \approx am^{b-1}$.

Le produit $Q(m) \cdot q(m) = H(m)$ définit un paramètre qui peut servir à l'estimation de l'énergie potentielle pour des individus de masses (m) variant de m_{min} à m_{max} intervalle dans lequel se situe la masse optimale m_{opt}, telle que $\log m_{opt} = \frac{1}{2} (\log m_{min} \times \log m_{max})$ ou $m_{opt} = (m_{min} \cdot m_{max})^{1/2}$.

L'optimum correspond donc à la moyenne géométrique des (m).

Le choix des paramètres α et β pour la transformation linéaire $\alpha\, q(m) + \beta = P(m)$ permet de définir

$$\begin{cases} P(m_{\min}) = Q(m_{\max}) \\ P(m_{\max}) = Q(m_{\min}) \end{cases} \quad \text{tel que} \quad P(m) = -\frac{dQ}{dq}\, q(m) + \frac{dH}{dq}$$

lorsque dQ, dq, dH représentent les augmentations de Q, q, et H dans l'étendue $(m_{\min})$ à $(m_{\max})$; pour $dH = 0$ on définit $m_0 = -\frac{dQ}{dq}$ ou en remplaçant Q et q par leurs expressions en $H(m)$ $Q(m) = [mH(m)]^{1/2}$ et $q(m) = [H(m)/m]^{1/2}$ d'où l'évaluation du point d'intersection $m_0 = (m_{\min} \cdot m_{\max})^{1/2}$ qui représente la masse optimale du point de vue énergétique.

Courbes de croissance intéressant des populations

Épidémiologie

a) Cas le plus simple : courbes de morbidité correspondant à cette seule alternative; sain ou malade : Exemple : Épidémie de fièvre typhoïde de Zermatt, 1963. Les fréquences de cas se déclarant à la fin de chaque demi-semaine depuis le début de l'épidémie (21 février) jusqu'à la fin (2 avril) se distribuent selon le modèle log normal (J. Dufrenoy, *Revue Path. Comp.*, n° 801, mars-avril 1969).

b) Modèle de la fonction logistique

La transformation en « logit » soit $\left(\log_e \frac{x}{1-x}\right)$ a été utilisée pour l'analyse statistique d'épidémies causées par des agents pathogènes dont le nombre augmente selon une fonction exponentielle : tel est le cas des épidémies (ou plus exactement épiphyties) dues aux « urédospores » des rouilles des graminées et des céréales. M. G. Eversmeyer et J. R. Burleigh (Phytopathology, 60, 805, 1970) transforment en logit la variable dépendante Y servant à estimer la gravité de l'épidémie, en fonction du logit de chacune des variables X_1, X_2..X_n; l'ordre dans lequel les variables indépendantes X_i sont introduites dans l'équation de régression est déterminé par l'estimation de F, qui permet de classer les X_i par ordre décroissant d'après la contribution de chaque X_i à la variation de Y dans l'équation de regression multiple de la forme $Y = b_1X_1 + \ldots + b_nX_n + K$.

Quant au développement d'une épidémie de rouille pour une culture donnée de blé, la 1re variable indépendante à introduire (X_1) correspond au nombre de urédospores dont dépend la gravité de l'exposition au risque d'infection; cette variable peut s'estimer en fonction du nombre des urédospores dans un volume donné de l'atmosphère, à chaque semaine successive, ou par la sommation des résultats des dénombrements hebdomadaires durant la période possible d'infection, ces nombres étant transformés en logit; la 2e variable à introduire rend compte des facteurs météorologiques : moyennes des minima et des maxima de température, nombre de jours de pluie, nombre d'heures par jour où les feuilles restent humides (rosée).

c) Modèle de la chaîne de MARKOV applicable à l'évaluation du risque de pneumoconiose.

Au cours de 35 années d'exposition au risque d'inhalation de poussières dans les mines de charbon, le mineur, initialement au stade 0 de la classification radiologique de « pneumoconiose » peut passer aux stades 1, 2, 3.

Chaque stade comporte 3 « sous stades » soit O/—; 0/0; 0/1; 1/0; 1/1; 1/2; 2/1; 2/2; 2/3; 3/2; 3/3; 3/4;

D'après les résultats d'examens radioscopiques de 4 122 mineurs dans 20 houillères de Grande-Bretagne, on a estimé la corrélation entre dose X mg de poussière par m³ d'air respirable ($x = 1$ à 8) et fréquences de passages d'un sous-stade au sous-stade supérieur; pour chaque houillère, on établit une « matrice de transition », dont les éléments, traités selon le modèle de la Chaîne de MARKOV, permettent d'estimer la probabilité (p) pour un mineur sans pneumoconiose (sous stade 0/0) de passer dans la catégorie 1/0.

Après transformation arc sin, les probabilités portées sur échelle arithmétique en ordonnées, chacune à son rang x mg/m³ de poussière respirable, définissent une courbe d'allure exponentielle, partant de $p = 0$ pour $x = 1{,}2$ et tendant vers $p = 50$ % pour $x = 8.2$

Les poussières des houillères étudiées contiennent 0,8 à 7,8 % de quartz qui ne paraît pas avoir un effet statistiquement significatif.

(M. Jacobsen, S. Rae, W. H. Walton & J. M. Rogan, Institute of Occupational Medicine, Edimbourg EH8, 95 U. *Nature*, 227, aug. 1, 1970, 445-447).

Modification des courbes de croissance d'une population par interférence avec la nutrition

Cinétique du développement d'une population de Tetranychus urticae sur feuilles de haricots traitées par chelates de métaux.

W. N. Cannon, *Jour. Econ. Entom.*, June 1970, p. 724-725.

Dans des conditions permettant une pullulation illimitée, le taux de croissance de la population peut s'estimer par l'équation.

$N_t = N_o\, e^r m^t$, lorsque N_o et N_t sont les nombres d'individus aux temps zéro et t, r_m, capacité intrinsèque d'accroissement, correspondant au taux instantané d'accroissement d'une population ayant atteint la stabilité quant à la distribution des âges, le milieu étant optimal, notamment quant aux facteurs de nutrition; x représentant la position centrale de chaque classe d'âge, l_x la proportion des naissances donnant des individus encore vivants dans chaque classe x, et m_x le nombre de naissance de femelles (ou le nombre des œufs produisant les femelles) par femelles dans chaque classe d'âge x; on pose $\Sigma\, e^r m^x l_x m_x = 1$, et on effectue des approximations successives en introduisant diverses valeurs de r.

Ayant évalué r_m pour une population de *Tetranychus urticae* sur feuille de haricot non traitée, on évalue r_m pour une population nourrie sur feuille traitée par Fe DTPA (penta acetate sodique ferrique de diethylenetriamine ou Mg — EDTA (tetraacetate disodique, magnesien, d'éthylène diamine).

Pour chaque traitement à chaque phase du cycle de développement on calcule : le taux net de reproduction $R_o = \Sigma\, l_x m'_x$ (pour chaque classe x); la capacité intrinsèque r_m; la durée moyenne d'une génération $T = ln\ R_o/r_m$; le taux final d'accroissement $\lambda = e^r m$; le temps au cours duquel la population double, d'après l'équation $N_t = N_o e^r m^t$.

Le cycle de développement débute lors de l'infestation d'une plante de haricot par une femelle adulte; les femelles, issues de ponte, sont dès l'âge adulte placées sur feuille non traitée ou feuille traitée par solution à 1 % Fe DTPA ou Mg EDTA.

Sur feuille non traitée, $R_o = 38.6$ (la population a l'aptitude de se multiplier 38,6 fois par génération); sur feuille Fe DTPA, $R_o = 22.8$; sur feuille Mg EDTA, $R_o = 19.9$; les

valeurs de r_m (déterminées pour une large part durant la première semaine de la période de reproduction) tombent de 0.23 sur feuille non traitée à 0.184 sur feuille traitée; la durée moyenne d'une génération T passe de 16 jours sur feuille non traitée à 17 sur feuille traitée par Fe DTPA; il suffit sur feuille non traitée de 3 jours pour que double la population; il en faut près de 4 sur feuille traitée.

Courbes de croissance de l'usage de la cigarette aux États-Unis de 1955 a 1966

d'après P. P. Ahmed & G. Gleeson, Nat. Center for Health Statistics, Séries 10, n° 59, avril avril 1970, pp. 1-33)

1) L'enquête par sondage, a porté sur 35 000 foyers, comprenant 68 000 personnes de plus de 18 ans, en 357 sites d'enquête, quant à l'alternative « non fumeur » (n'ayant jamais fumé ou ayant cessé de fumer) ou utilisant par jour plus d'1 cigarette (soit 1 à 9 ou 10 à 20, ou 21 à 40, ou plus de 40).

Les résultats sont affectés :

a) par les réponses erronées;

b) par les erreurs des enquêteurs;

c) par la variabilité inhérente à l'échantillonnage, et qui s'estime par la déviation standard d'échantillonnage (quant aux variations aléatoires imputables au fait qu'on n'a étudié qu'un échantillon tiré de la population, au lieu de l'ensemble de la population); il y a 68 chances sur 100 pour que l'estimation d'après l'échantillon diffère d'avec l'estimation de l'ensemble de la population, de moins de 1 déviation standard (p. 29).

Par exemple, si le nombre total des fumeurs de (*c*) cigarettes a été estimé à 750 000 en 1966 (avec erreur standard 35 000), à 600 000 en 1955) avec erreur standard 41 000), l'erreur standard de la différence entre ces deux estimations de dex caractéristiques supposées non liées par corrélation, est estimée par la racine carrée de la somme des carrés des 2 déviations standard, soit ici $[(35)\ (10^3) + (41)\ (10^3)]^{1/2} \approx (54)\ (10^3)$; il y aurait 68 chances sur 100 pour que la différence estimée n'atteigne pas $(54)\ (10^3)$ si l'hypothèse zéro est vraie (s'il n'y a pas de différence réelle entre les résultats de 1966 et 1955); en prenant comme seuil la probabilité $p = 0.95$, de considérer comme réelle la différence correspondant à 2 déviations standard, soit $(108)\ (10^3)$, on prend 5 % de chances de considérer à tort, comme réelle, la différence entre 750 000 et 600 000 (p. 29).

Analyse statistique des résultats :

1) Les résultats étant publiés sous forme de pourcentages, les comparaisons peuvent s'effectuer facilement par le calcul du critérium de Chi², pour chaque table de contingence utilisant deux sources d'information, par exemple classification par colonne, 1955 ou 1966, classification par rang, selon année à laquelle a été acquise l'habitude de fumer : voir exemple numérique (p. 6).

2) Le degré de confiance à accorder aux résultats publiés peut être estimé d'après la conformité de distribution des pourcentages « observés » avec tel ou tel modèle mathématique de distribution — voir page 5, modèle mathématique utilisant la transformation doublement logarithmique — $ln\ (-\ ln\ n)$ des pourcentages cumulés n (graphique 18).

Augmentation des usagers de cigarette dès l'age de 10 ans

En juin 1966, pour « les plus de 18 ans », 49,3 % des 56 768 000 hommes (27 998 000) et 32,1 % des 64 310 000 femmes (20 670 000) fumaient plus d'une cigarette par jour; en 1955 les pourcentages étaient 53,8 et 27,4;

La population mâle de plus de 18 ans ayant augmenté de 7 187 000, malgré la diminution du pourcentage, le nombre des fumeurs a augmenté de 1 300 000 (p. 5).

L'augmentation a été relativement plus considérable chez les femmes; à l'accroissement de 7 187 000 femmes correspond un accroissement de 7 187 000 femmes qui fument. Alors qu'en 1955, moins d'une femme sur mille avait commencé à fumer à 14 ans en 1966, 1,4 % avaient acquis l'habitude dès l'âge de 10 ans (table 8, p. 23).

Les pourcentages cumulés de femmes fumant depuis l'âge de 10, 11 années, atteignant en 1966, les totaux de 44,9 % pour les femmes de 18 à 24 ans; 50,3 % pour celles de 25 à 34 ans.

Au lieu des pourcentages cumulés de femmes fumant la cigarette, on peut calculer les pourcentages cumulés de femmes qui aux âges 10, 11 ... 24 ans, n'avaient pas contracté l'habitude de fumer : en portant ces pourcentages (n) de non fumeurs à âge x pour $x = 10$ à $x = 24$, en ordonnées sur échelle $-\ln(-\ln n)$ chacun à son rang d'abscisse x, on détermine, pour 1966, dix points par lesquels on peut tracer une droite de régression, partant du niveau 98,6 % à $x = 10$ et atteignant le niveau 65 % à $x = 19$, pour 1966; comparativement pour 1955 on détermine, à partir de $n = 99{,}1$ à $x = 14$, jusqu'à $n = 76$ à $x = 19$, une droite de régression de pente plus accusée que celle de la droite 1966.

Le graphique suggère que la diminution des pourcentages de femmes n'ayant pas acquis à l'âge $x = 10$ à $x = 14$ l'habitude de fumer la cigarette est conforme à la distribution log. normale qui caractérise un phénomène contagieux : la comparaison des distributions pour 1955 et 1966 suggère qu'il s'agit d'un processus Markovien, le pourcentage des femmes ayant déjà acquis à l'âge x l'habitude de fumer dépend du pourcentage à l'âge $x - 1$.

C'est ce que suggère la table de contingence à 2 classifications, par colonne, 1955 et 1966, et par rang selon âge x d'acquisition de l'habitude de fumer.

Pour les 4 classes d'âge $x = 14$ à $x = 17$, on calcule $2\hat{i} = 10$, qui pour 3 degrés de liberté à moins de 0.015 % de chance de se manifester du seul fait du hasard. On peut donc considérer comme significative l'augmentation des pourcentages de 1955 à 1966 dans ces tranches d'âge.

Tableau C

c	Hommes			Femmes		
		Cumul	n		Cumul	n
0	49.1			65.7		
1 à 10 . . .	5.6	5.6	11.2	5.7	5.7	17.1
11 à 20 . . .	28.9	34.5	69.0	22.1	27.8	83.4
21 à 40 . . .	13.1	47.6	95.2	5.0	32.8	98.4
plus de 40 . .	1.8	49.4	98.8	0.3	33.1	99.3

Du tableau C nous tirons les pourcentages par classes pour 1966; ne considérant que les usagers de plus d'une cigarette par jour, nous cumulons les pourcentages par classe 1 à 10; 11 à 20; 21 à 40; plus de 40; nous calculons ces pourcentages en fonction du total et nous portons ces pourcentages (n) en ordonnée sur échelle — ln (— ln n) chacun à son rang de limite de classe, 10, 20, 40, sur échelle log en abscisses. On définit pour les hommes et pour les femmes, 3 points parmi lesquels on trace une droite de régression; pour les femmes, cette droite coupe le niveau 50 % à l'aplomb de $C = 13$; le niveau 84 % à l'aplomb de $C = 20$, ce qui suggère que 50 % forment au plus 13 cigarettes, 84 % au plus 20; pour les hommes la droite coupe le niveau 50 % à l'aplomb de $c = 15.5$, le niveau 84 % à l'aplomb de $C = 28$, ce qui suggère que 50 % des fumeurs fument moins de 16 cigarettes, 84 % au plus 28.

Ces valeurs de $Q_{50} = 15.5$ pour les fumeurs de 13 pour les femmes, représentent pour (28) $(10)^6$ fumeurs 434 millions de cigarettes par jour et pour (20.5) (10^6) fumant la cigarette 260 millions, soit près de 700 millions au total.

En attribuant à chaque classe la valeur intermédiaire entre les 2 limites inférieure et supérieure on parvient à des estimations de 544 millions pour les fumeurs, 346 pour les femmes (table 10, p. 26) soit près de 1 milliard de cigarettes fumées par jour aux États-Unis.

FRÉQUENCES PAR CLASSES EN MILLIONS

Cigarettes par jour	Hommes	Femmes		Cigarettes en millions	
				Hommes	Femmes
1 à 10	1.2	2.3	5	6	10
11 à 20	13.3	12.1	15.5	207	182
21 à 40	8.8	4.5	30	264	132
plus de 40	1.5	0.5	45	67	22
				544	346

(En 1969 le S. E. I. T. A. avait fabriqué en France 70 554 millions de cigarettes).

L'échelle log des abscisses ne permet pas de faire figurer les « non-fumeurs » ($c = 0$); l'échelle arc-tangente admet comme origine zéro; à partir de cette origine les valeurs telles que $c = 10$, 20 ou 40 peuvent être espacées à peu près comme sur une échelle log.

Reprenant les pourcentages par classe $c = 0$ à $c > 40$, cumulant ces pourcentages, et portant les valeurs ainsi obtenues de n sur échelle (— ln — ln n) chacune à son rang $c = 0$, $c = 10$, $c = 20$ ou $c = 40$ sur échelle arc tangente on définit encore pour $c = 10$, $c = 20$, $c = 40$, trois points par lesquels on peut tracer une droite de régression;

84 % de l'ensemble de la population masculine, avec environ 50 % de non-fumeurs et 50 % de fumeurs, consomment par individu au plus 20 cigarettes et par conséquent 16 % plus de 20 cigarettes.

Pour l'ensemble de la population féminine avec seulement 33 % d'usagers de cigarettes, 84 % consomment moins de 15 cigarettes.

Effets comparés de la sous-alimentation et de la déficience en protéine sur les courbes de croissance du rat

(S. Reindorf and R. G. White head *Nature*, 227, 404, 6, 70)

Dès le sevrage, au 28e jour, des rats (var. black-hooded) consomment à volonté une ration :

a) à 16 % de protéine,

b) à 6 % de protéine :

Comparativement les rats du lot *C* reçoivent la même ration que (*a*) mais en quantité assez réduite pour que leur croissance soit ralentie comme celle de (*b*).

5 rats sont sacrifiés après x jours ($x = 0, 3, 7, 15, 30, 60$) et on détermine les poids (y) du corps, du cœur, du muscle « tibialis » et du foie.

Après transformation des $x = 0, 3, 7, 15$ en coefficients polynómiaux Z_1, Z_2, Z_3, nous avons calculé les équations polynomiales $\hat{y} = \bar{y} + BZ_1 + CZ_2 + DZ_3$ comparativement pour *a*, *b*, *c*.

1 — *Poids du corps y* g

Tableau I

x	Z_1	Z_2	Z_3	a	b	c
0	— 7	+ 7	— 3	43	43	43
15	— 3	— 4	+ 8	116	58	59
30	1	— 8	— 6	171	82	82
60	9	5	1	244	96	95
	140	154	110	574	279	279

Exemple numérique : Rats (*a*)

x	Z_1y	Z_2y	$\hat{y} = 143.5 + 12{,}264\, Z_1 - 2.02\, Z_2$	$(y - \bar{y})$	$(y - \bar{y})^2$
0	— 301	301	143.5 + [(— 7) (12.271) + (7) (— 2.02)] = 43.130	— 100.5	10 100,25
15	— 348	— 464	143.5 + [(— 3) (12.271) + (— 4) (— 2.02)] = 114.767	— 27.5	756.25
30	171	— 1368	143.5 + [(1) (12.271) + (— 8) (— 2.02)] = 171.931	+ 27.5	756.25
60	2196	1220	143.5 + [(9) (12.271) + (5) (— 2.02)] = 243.849	+ 100.5	10 100.25
	1718	— 311	573.679		21 713

Analyse de la variance

Source de variation	Somme des carrés	*DL*	*F*
Totale $\Sigma\,(y - \bar{y})^2$	21 713		
Due à régression :			$\frac{21080.48}{2.62} = 8046$
linéaire B [$\Sigma\,(Z_1\,y)$]	21 080,48	1	
Quad. C [$\Sigma\,(Z_2\,y)$]	626.04	1	$\frac{626.04}{2.62} = 240$
Cub. D [$\Sigma\,(Z_3\,y)$]	2.62	1	

La variance imputable à la régression quadratique $D\,[\Sigma\,(Z_3 y)] = (17)\ (0.154) = 2.62$ correspond (aux approximations près) à la variance erreur $\Sigma\,(y - \hat{y})^2 = 2.2914$. Les valeurs de F, pour 1 DL au numérateur et 1 DL au numérateur, indiquent que les variances dues à régression linéaire, puis à régression quadratique sont significatives; après transformation en Z_1 et Z_2 la courbe de croissance se représente pour les rats (a) par un arc de parabole passant par le point $(\bar{x}, \bar{y})$ avec $\bar{x} = 26.25$ et $\bar{y} = 143.5$; les déviations $(y - \bar{y})$ étant symétriques, soit -100.5 pour $x = 0$, et $+100{,}5$ pour $x = 60$; puis -756.25 pour $x = 15$ et $+756.25$ pour $x = 30$.

La comparaison des équations polynomiales pour les rats (a) (b) et (c) met en évidence la modification de la courbe de croissance sous l'effet de la malnutrition (b) ou sous nutrition (c), par diminution de la moyenne $(\bar{y})$, des coefficients (B) et (C) de régressions linéaires et quadratique, tandis qu'augmente le coefficient (D) de régression cubique :

$$a)\ \hat{y} = 143.5 + 12.264\,Z_1 - 2.02\,Z_2 + 0.008\,Z_3$$
$$b)\ \hat{y} = 69.75 + 3.364\,Z_1 - 0.695\,Z_2 - 0.55\,Z_3$$
$$c)\ \hat{y} = 69.75 + 3.27\,Z_1 - 0.753\,Z_2 - 0.49\,Z_3$$

2 — *Allométrie*

Comparablement aux équations ci-dessus d'accroissements de poids du corps, nous calculons les équations pour les accroissements de poids de muscles du squelette *(tibialis)* ou du cœur, ainsi que du foie.

TABLEAU II

x	Muscles (y mg)						Foie y g		
	tibialis			cardiaque					
	a	b	c	a	b	c	a	b	c
0	62	62	62	231	231	231	1.70	1 70	1.70
15	189	101	103	457	249	287	5 36	3 04	2 40
30	282	139	147	613	347	296	7 08	3 69	2 56
60	460	198	208	701	344	344	8 92	3 93	3.27
	993	500	520	2002	1171	1098	24 04	12 36	9.93

Muscle tibialis (y mg)

$$a)\ \hat{y} = 248.25 + 24.4\,Z_1 + 3.4\,Z_2 + 0.85\,Z_3$$
$$b)\ \hat{y} = 125 + 8.45\,Z_1 - 0.62\,Z_2 + 0.127\,Z_3$$
$$c)\ \hat{y} = 130 + 8.22\,Z_1 - 0.74\,Z_2 - 0.32\,Z_3$$

Muscle cardiaque (y mg)

$$a)\ \hat{y} = 500.5 + 28.1\,Z_1 - 10.4\,Z_2 - 0.15\,Z_3$$
$$b)\ \hat{y} = 292.75 + 7.70\,Z_1 - 2.82\,Z_2 - 4\,Z_3$$
$$c)\ \hat{y} = 274.45 + 6.35\,Z_1 + 0.81\,Z_2 - 3\,Z_3$$

Foie (y g)

$a)\ \hat{y} = 6.01 + 0.5\ Z_1 - 0.0752\ Z_2 + 0.038\ Z_3$

$b)\ \hat{y} = 3.09 + 0.13\ Z_1 - 0.0655\ Z_2 + 0.0007\ Z_3$

$c)\ \hat{y} = 2.4825 + 0.092\ Z_1 - 0.0134\ Z_2 - 0.0364\ Z_3$

3 — *Courbes d'accroissement de la taille*

Les longueurs (y cm) mesurées de la pointe du museau à la naissance de la queue, connaissent de $x = 0$ à $x = 60$, un accroissement moindre pour (b) et (c) que pour (a) mais les équations polynomiales révèlent une différence beaucoup moins importante que par les accroissements pondéraux (tableau III).

TABLEAU III

x	a	b	c
0	12 6	12 6	12 6
15	16 4	13 8	14 0
30	19.1	16.3	15.9
60	21.4	16.0	16 4
	69.5	58.7	58 9

$a)\ \hat{y} = 17.375 + 0.53\ Z_1 - 0.15\ Z_2 + 0.00018\ Z_3$

$b)\ \hat{y} = 14.675 + 0.30\ Z_1 - 0.12\ Z_2 + 0.146\ Z_3$

$c)\ \hat{y} = 14.75 + 0.39\ Z_1 - 0.081\ Z_2 - 0.05\ Z_3$

Les rats (a) atteignent en 15 jours la longueur de 16.4 cm que les rats (c) n'atteignent qu'en 60 jours.

Pour $x = 0, 3, 7, 15$, valeurs qui peuvent approximativement être considérées comme équivalentes à 0, 1, 2, 4, on calcule $y = 14.2 + 0.246\ Z_1 - 0.035\ Z_2 - 0.2\ Z_3$ équation comparable à celle calculée de $x = 0$ à $x = 60$ pour (c).

4 — *Accroissement pondéral en fonction de la taille*

Muscle cardiaque (y mg/10 cm).

TABLEAU IV *a*

x	a	b	c
0	182.9	182.9	182.9
16	278.2	179.8	162.0
30	332.4	212.7	187.5
60	327.7	214 4	209 4
	1121 2	789.8	741.8

$a)\ \hat{y} = 280.3 + 8.334\ Z_1 - 5.47\ Z_2 + 0.0927\ Z_3$

$b)\ \hat{y} = 197.45 + 2.8\ Z_1 - 1.4\ Z_2 - 1.8\ Z_3$

$c)\ \hat{y} = 1845.45 + 1.123\ Z_1 - 0.65\ Z_2 + 1.51\ Z_3$

Pour les rats (*a*) l'équation polynomiale pour (y mg/10 cm) est, comme celle de l'accroissement de poids du corps, l'équation d'un arc de parabole, le coefficient de régression cubique (en Z_3) pouvant être négligé; pour les rats (*b*) ou encore plus pour les rats (*c*) les poids du cœur exprimés en fonction de la taille n'augmentent pas significativement (le coefficient de régression linéaire n'étant que 2.8 pour *b* et 1.123 pour *c*, au lieu de 8.334 pour *a*).

Poids du foie (y g/10 cm)

TABLEAU IV *b*

	a	*b*	*c*
0	1.35	1.35	1.35
15	3.27	2.21	1.71
30	3 97	2.38	1.59
60	4.18	2.45	1.98
	12.77	8.39	6.63

$$a)\ y = 3.1925 + 0.1595\ Z_1 - 0.09\ \ Z_2 + 0.03\ \ Z_3$$
$$b)\ \hat{y} = 2.0975 + 0.0596\ Z_1 - 0.04\ \ Z_2 + 0.016\ Z_3$$
$$c)\ \hat{y} = 1.6575 + 0.0331\ Z_1 - 0.00007\ Z_2 - 0.02\ \ Z_3$$

Les équations polynomiales sont conformes aux conclusions de S. R. et R. G. W. : en fonction de la taille y cm, les foies augmentent de poids, des jours $x = 0$ à $x = 60$, pour (*c*) et encore plus nettement pour (*b*), la déficience de protéine provoquant l'infiltration de graisses;

5 — *Effets spécifiques de la déficience protéique ou de la sous-nutrition sur le cœur*

S. R. et R. G. écrivent « Encore plus manifestement que le muscle squelettique (tibialis) le muscle cardiaque est affecté par sous-nutrition : le poids du cœur augmente régulièrement et triple en 60 jours, chez les témoins, alors qu'il n'augmente que de 24 % pour (*c*) ou (*b*).

L'équation polynomiale pour (*a*) est celle d'une courbe parabolique (accroissement pondéral moins que proportionnel au temps). Pour (*b*) et (*c*) le coefficient de régression linéaire (en Z_1) correspond à 25 % du coefficient pour (*a*), tandis que le coefficient de régression cubique (terme en Z_3) est relativement important. La perturbation de la croissance du muscle cardiaque peut être corrélative d'une réduction de la circulation sanguine dans un organisme émacié, la croissance pondérale étant relativement plus gravement ralenti que la croissance en longueur.

Une réduction relative du poids du cœur s'observe chez les enfants souffrant du Kwashiorkor, en Ouganda, les accidents cardiaques, corrélatifs d'anémie et de taux excessif de Na dans le sang, constituent un risque grave au cours du traitement; des complications cardiaques ont été signalées au Mexique, à la Jamaïque et au Guatemala.

Conclusions

Pour chaque régime *a*, *b*, *c*, 6 pesées ou mensurations effectuées aux jours $x = 0$, 3, 7, 15, 30, 60, ont permis de déterminer :

1° des poids (y g) et des longueurs (y cm) du corps,

2° des poids (y mg) de muscles tibialis ou cardiaque et (y g) de foie,

3° d'estimer chaque poids (y g) ou (y mg) par 10 cm de longueur de l'animal : la croissance de (y) en fonction du temps (x) se représente théoriquement par une courbe en *S* (dite sigmoïde) qui peut être calculée par une équation du 3e degré en x; par 4 points, on peut toujours faire passer une courbe du 3e degré en x; une courbe sigmoïde, peut être définie par 4 points ($x_i y_i$); les calculs sont particulièrement faciles si les x correspondant à une série telle que 0, 1, 2, 4; chaque valeur de x peut alors être remplacée par $Z_1 = \lambda (x - \bar{x})$ avec $\bar{x} = \frac{0 + 1 + 2 + 4}{4} = \frac{7}{4} = 1.75$. On fait correspondre à $x = 0$, 1, 2, 4, les valeurs de Z_1 telles que $4\ (0 - 1.75) = -7$; $4\ (1 - 1.75) = -3$; $4\ (2 - 1.75) = 1$ et $4\ (4 - 1.75) = 9$.

Ces valeurs de Z_1 sont orthogonales $(-7 - 3 + 1 + 9) = 0$.

On transforme de même les x^2 en Z_2, et les x^3 en Z_3.

L'équation polynomiale s'écrit alors $\hat{y} = \bar{y} + BZ_1 + CZ_2 + DZ_3$; l'analyse de la variance permet d'attribuer successivement à *B*, à *C* puis à *D* sa contribution à la variance totale et, si elle n'est pas significative, de l'imputer à la variance erreur.

Le grand intérêt de la série $x = 0$, 1, 2, 4 apparaît notamment par comparaison avec une série telle que 1, 2, 3, 4 étudiée ci-après.

Effets de la teneur des protéines en amino-acides sur la courbe de croissance des rats.

J. Abraham, E. Baratou, P. Robin et R. Jacquot (*Ann. Nutr. Alim.*, 1970, 24, 51-91) nourrissent comparativement, après sevrage des rats avec grains de maïs commercial (protéine 2.4 lysine) ou de opaque-2 (protéine à 4,8 % lysine) observent les consommations y_c g/jour et les gains de poids y_p g/jour au bout de $x = 1$, 2, 3, 4 semaines, avec maïs.

x	Commercial			Opaque 2		
	y_c	y_p	y_p	y_c	y_p	$\hat{y}_p$
1	74.4	8.6	8.610	93.5	22.6	22.63
2	82.7	12.2	12.155	115 4	26.5	26.41
3	90.9	13.7	13 745	127.1	27 4	27.49
4	90.2	13.4	13.385	128.0	25.9	25.87
	388.0	47.9		464.0	102.4	

Maïs commercial	$\hat{y}_c = 84.55 + 2.78\ Z_1 - 2.25\ Z_2$	$\hat{y}_p = 11.975 + 0.795\ Z_1 - 0.975\ Z_2$
Opaque 2	$\hat{y}_c = 116 + 5.76\ Z_1 - 3.5\ Z_2$	$\hat{y}_p = 25.6 + 0.54\ Z_1 - 1.35\ Z_2$

Le temps pris comme variable dépendante

Les équations de courbes de croissance s'expriment généralement sous la forme $Y = f(X)$, la variable indépendante X représentant le temps, porté en abscisses sur échelle appropriée; la variable dépendante (Y) prenant pour chaque valeur (x) attribuée à (X) une valeur (y) exprimée en unités de poids ou de longueur.

Étudiant « l'influence de la vitamine A sur réplétion protéique du rat blanc après inanition azotée prolongée » (*Ann. Nutr. Alim.*, 1970, 24, 93-105) S. Blaizot, J. Blaizot et M. Caussanel remarquent que les différences de poids entre témoins et carencés expriment une disparité de vitesses de croissance : après une durée (x) fixée *a priori*, de régime de carence, les poids peuvent différer considérablement entre carencés et témoins; pour comparer témoins et carencés, les auteurs ne fixent pas une durée (t) de carence, mais un poids (P) à atteindre par les animaux.

Exemple : 42 rats sont « préparés » par deux semaines de régime à 18 % de caséine sans vitamine A, puis chacun des 14 rats de chacun des 3 lots reçoit 0.20 UI ou 200 UI de vitamine pendant le nombre de jours requis pour acquérir le poids $P_0 = 150$ g.

La caséine est alors remplacée par des glucides. Chaque rat est privé d'aliment azoté pendant le nombre de jours j nécessaire pour une perte de 1/3 P_0; le poids tombant à 150 — 50 = 100 g. Ce temps j peut être divisé en 3 périodes j_1, j_2, j_3 chacune correspondant à la perte de 1/9 de P_0; la variable indépendante prend donc les valeurs successives 1/9 P_0, 2/9 P_0, 3/9 P_0 formant les termes d'une progression arithmétique; la variable dépendante (Y) prend alors les valeurs y indiquées dans le tableau et utilisées pour le calcul des équations polynomiales $\hat{y} = \bar{y} + BZ_1 + CZ_2$ pour chaque traitement 0, ou 20 ou 200 UI, vitamine A.

Vitamine A	x			Total
	1	2	3	
0	6.9	13.2	19.3	39.4
20	6.4	15.8	24.7	46.9
200	6.6	16.1	22.9	45.6

$$y = 13.133 + 6.2\ Z_1 - 0.033\ Z_2$$
$$\hat{y} = 15.633 + 9.15\ Z_1 - 0.083\ Z_2$$
$$y = 15.2 + 8.15\ Z_1 - 0.35\ Z_2$$

Analyse de la variance

Vitamine A		0	20	200
Source de variation	*DL*	*SC*	*SC*	*SC*
Totale $\Sigma\ (y - \bar{y})^2$	1	76.9518	167.4578	134.06
due à régression :				
linéaire $B\ \Sigma\ (Z_1 y)$	1	76.880	167.445	132.845
quad $C\ \Sigma\ (Z_2 y)$	1	0.0066	0.415	0.785

Selon les auteurs : « la chute de poids est très rapide au début de la déplétion protéique et évolue ensuite plus lentement; pour passer de 150 à 100 g les rats carencés emploient moins de 40 jours, contre plus de 45 pour les rats qui ont reçu de la vitamine A ». L'équation polynomiale pour les rats carencés (0 vitamine A) diffère les deux autres par les faibles valeurs de la moyenne et des coefficients de régressions linéaire et quadratique : la valeur du coefficient de régression quadratique (en Z_2) augmente de 0 à 20 puis 200 UI, vitamine A.

Calcul des équations de courbes de croissance par polynomes orthogonaux

Sous le titre « Analyse des expériences effectuées en vue d'estimer des taux de croissance », F. B. Leech et M. J. R. Healy (*Biometrics*, 15, 1er mars 1959) envisagent une expérience au cours de laquelle l'un des divers traitements est appliqué à chaque groupe d'individus, au cours du cycle de croissance; la variable dépendante Y prend successivement les valeurs y_{t_0}, y_{t_1} ... y_{t_n} aux époques t_0, t_1 ... t_n. Du début (t_0) de l'expérience à la fin (t_n), la croissance totale a été $y_{t_n} - y_{t_0}$; le quotient $\frac{(y_{t_n} - y_{t_0})}{(t_n - t_0)}$ fournit une évaluation du taux moyen de croissance; cette évaluation n'utilise pas l'information acquise au cours de chaque période successive t_1, t_2 . . . $t_{(n-1)}$, par la mesure des valeurs y_{t_1}, y_{t_i} . . . $y_{t_{(n-1)}}$; l'analyse statistique de l'ensemble des résultats de mesures doit tenir compte de ce que les résultats successifs tels que y_{t_i} dépend des valeurs antécédentes y_{t_0}, y_{t_1} . . . $y_{t_{(i-1)}}$.

Dès 1938-39, J. Wishart (Biométrika, 30, 16-28; *J. Roy, Stat. Soc. Suppl.*, 6, 1-22) avait suggéré de décomposer l'équation de la courbe de croissance, quant à ses différents paramètres : moyenne $\sum_1^n \frac{(y_i)}{n} = \bar{y}$ pour n observations et coefficients de régression linéaire, (tendance à accroissement proportionnel au temps), quadratique (tendance à accroissement plus que proportionnel ou moins que proportionnel), cubique...

Les effets de chaque « traitement » sur l'allure de la courbe de croissance pevent alors être estimés en fonction de la valeur de chaque paramètre, qui peut être pondéré par l'estimation des variances et de la covariance.

L'analyse peut facilement s'effectuer en calculant pour les résultats (y_t) des n observations chacune effectuée au temps t, les polynomes orthogonaux jusqu'au degré (n — 1) en t, lorsque t représente la variable indépendante, généralement désignée par x; l'équation polynomiale a la forme

$$\hat{y} = \bar{y} + BZ_1 + CZ_2 + \dots KZ_i \dots + NZ_{(n-1)}$$

lorsque Z_i représente la valeur numérique polynomiale de degré i pour la valeur x correspondant au temps t_i de l'observation y_t dont on estime la valeur par $\hat{y}_t$.

Pour une population d'individus recevant tous le même traitement, les valeurs des paramètres $\bar{y}_1$, B_1, C_1... peuvent être considérées comme les valeurs estimées de variables aléatoires.

Pour une 2e population soumise à un traitement différent, on peut de même estimer les paramètres $\bar{y}_2$, B_2, C_2...

Cas simple de blocs randomisés chacun comportant 2 sujets : « témoin » T, et « expérimental » E; on estime la différence entre chaque paire successive T et E à chaque époque t_i et on calcule l'équation polynomiale pour les différences $(y_T - y_E) = d_i$.

Dans l'hypothèse (H_0) où l'effet du traitement est proportionnel au temps, à la différence entre les courbes de croissance des populations 1 et 2, correspond une ligne droite passant par l'origine et de pente B_1 à estimer.

L'hypothèse (H_0) peut aussi s'exprimer ainsi : le traitement n'a pas d'effet sur les composantes quadratiques (C), cubique (D)... On critique H_0 en effectuant l'analyse de la variance permettant d'estimer la contribution de chaque régression : linéaire, quadratique, cubique... à la variance totale.

Exemple numérique : cinq observations successives $(n = 5)$ effectuées à l'origine $(t = 0)$ puis à $t = 1, 2, 3, 4$, c'est-à-dire à intervalles également espacés (pouvant correspondre à n'importe quelle unité de mesure du temps); le tableau indique les valeurs :

1° des coefficients polynomiaux; 2° de la somme des carrés de leurs valeurs.

Enfin, en vue d'estimer des polynomiales de degré 1 et 2, passant par l'origine, on ajoute à chaque valeur de Z_0, Z_1 ou Z_2 une constante telle que $Z_0 = Z_1 = Z_2 = 0$ à $t = 0$, en vue d'estimer les valeurs de p ci-dessous; à chaque Z_0 on attribue la valeur 1; de telle sorte que la somme des Z_0, soit ici, $\sum_1^5 Z_0 = 5$; chaque p_{i0} correspond à n fois la valeur de $(-Z_i)$ à $t = 0$.

	Z_0	Z_1	Z_2	Z_3		$Ø_1$	$Ø_2$	$Ø_3$
0	1	— 2	2	— 1		0	0	0
1	1	— 1	— 1	2		1	—3	3
2	1	0	— 2	0		2	—4	1
3	1	1	— 1	— 2		3	—3	—1
4	1	2	2	1		4	0	2
K	5	10	14	10				

$$p_{11} = 10, \qquad p_{10} = 10$$
$$p_{22} = 14, \qquad p_{20} = -10$$
$$p_{33} = 10, \qquad p_{30} = 5$$

L'effet du traitement étant quadratique en fonction du temps, l'équation polynomiale comporte un terme en Z_2 ou son équivalent en Ø 2.

On peut alors obtenir ces deux estimations de la pente correspondant à la régression linéaire.

$$l_0 = \frac{\left(k_0 d_0 - \frac{p_{20}}{p_{22}} K_2 d_2\right)}{p_{10}} \quad \text{et} \quad l_1 = \frac{K_1 d_1}{p_{11}}$$

pour $n = 5$ on a $K_0 = 5$; $K_1 = 10$ et $K_2 = 14$.

Effets sur la croissance des rats albinos mâles, d'ingestion de sulforam d'après H. G. Holck et al. *Jour. Pharmac. Sci.*, septembre 1970, pp. 1267-1270.

Le tableau indique pour deux séries d'expériences (I et II) les poids en grammes à l'âge x_j, d'après x_d jours de régime comportant 0, 1/2000, 1/1000 sulforam ou 1/1000 sulforam + 0,5 % ac. ascorbique.

x_j	x_d	0		1/2 000		1/1 000		1/1 000
		I	II	I	II	I	II	+ asc ac
30	0	55	56	56	56	56	56	54
60	30	148	149	129	131	115	122	85
90	60	191	193	169	175	157	161	148
120	90	223	222	195	192	185	184	200
		617	620	549	554	513	523	487
		154 25	155	137 25	138 25	128 75	130 75	121 75

Équations polynomiales

Sulforam	
0	I. $\hat{y} = 154\ 25 + 24{,}35\ Z_1 - 15\ 25\ Z_2$
	II. $\hat{y} = 155 + 27{,}01\ Z_1 - 16\ Z_2$
1/2 000	I. $\hat{y} = 137\ 25 + 27{,}85\ Z_1 - 11\ 75\ Z_2$
	II. $\hat{y} = 138\ 2 + 22\ 6\ Z_1 - 14\ 5\ Z_2$
1/1 000	I. $\hat{y} = 128{,}75 + 21\ 85\ Z_1 - 7\ 72\ Z_2$
	II. $\hat{y} = 130\ 75 + 21\ 15\ Z_1 - 10\ 75\ Z_2$
1/1 000 + ac asc	$\hat{y} = 121\ 75 + 20\ 5\ Z_1 - 5\ 25\ Z_2$

J. Dufrenoy