B. CYFFERS

Détermination du rendement du matériel de râpage du tabac à priser

Revue de statistique appliquée, tome 2, n° 3 (1954), p. 75-86
<http://www.numdam.org/item?id=RSA_1954__2_3_75_0>

DÉTERMINATION DU RENDEMENT DU MATÉRIEL DE RAPAGE DU TABAC A PRISER

par

B. CYFFERS

Ingénieur des Manufactures de l'État

Cette étude, entreprise en vue de déterminer le rendement réalisable par le matériel de râpage utilisé dans la fabrication de la poudre à priser est une application de la théorie de la régression. On y rencontre, tout d'abord, une régression multiple, puis une régression simple, et enfin une régression curviligne, traitée par la méthode des polynômes orthogonaux de Gauss.

Considérations sur la fabrication des tabacs à priser :

Afin d'éclairer le lecteur sur les quantités intervenant dans les calculs, il est indispensable de donner quelques précisions sur la fabrication de la poudre :

Les différents tabacs entrant dans la composition du produit sont mélangés, hachés à sec, et ensuite mouillés à l'eau salée chaude. On construit alors une masse contenant 65 tonnes de tabac environ que l'on soumet à fermentation durant 5 à 6 mois. On active ou ralentit cette fermentation, grâce à une aération permettant d'introduire une quantité d'air variable par le dessous de la masse. Des thermomètres, plantés à cœur de masse, en 9 points différents, permettent de suivre l'allure de la fermentation. La température s'élève aux environs de 85°, se stabilise, et décroît par la suite.

La masse est ensuite démolie et le tabac fermenté est livré à l'atelier du râpage. Le matériel de râpage est constitué essentiellement de moulins, analogues à de gros moulins à café, comprenant une noix tournant dans une cuvette.

Facteurs agissant sur le rendement du râpage :

On constate depuis longtemps que le rendement horaire moyen réalisé par moulin est très variable. Les facteurs ayant une action sur ce rendement sont certainement très nombreux. Nous n'avons retenu, pour cette étude, que ceux qui, selon les données de l'expérience et du bon sens, paraissent avoir une influence prépondérante. D'autre part, afin d'entreprendre les calculs à bref délai, nous avons dû choisir des quantités ayant fait dans le passé l'objet de mesures systématiques ; en effet, il est râpé environ 12 masses par an, et il aurait fallu prévoir un programme d'expérience sur plusieurs années pour disposer d'un nombre de données numériques suffisant.

Les facteurs retenus « à priori » sont :

— **la fermentation** : C'est un fait d'expérience que plus la fermentation a été active, plus le rendement des moulins est élevé. Comme « mesure » de la fermentation, nous avons adopté la moyenne arithmétique des températures maxima atteintes par chacun des 9 thermomètres. Soit T cette température moyenne.

— **l'humidité** : Il est logique de penser que plus le tabac est sec, plus il doit être facile de le réduire en poudre. Comme « mesure » de l'humidité du tabac, à son entrée au râpage, nous avons adopté l'excédent de mouillade, c'est-à-dire le poids d'eau salée ajouté à 100 kgs de tabac lors de la construction des masses. Soit M cet excédent de mouillade. M ne représente certainement qu'assez approximativement l'humidité réelle du tabac lors de sa livraison au râpage.

Enfin, nous désignerons par R le rendement horaire moyen d'un moulin.

Afin d'éviter de manipuler, dans les calculs qui vont suivre, des nombres trop grands, nous avons adopté les origines suivantes pour les mesures de R, M, T :

R : 21 kg/heure ;

M : 5 (M est un pourcentage) ;

T : 67°.

Pour bien fixer les idées, décidons de désigner par R, M, T, les valeurs du rendement, de l'excédent de mouillade, de la température, comptées à partir de ces origines, et par R_o, M_o, T_o les valeurs **réelles** de ces mêmes quantités.

Détermination du rendement en fonction de la fermentation et de l'humidité :

Les données numériques utilisées sont relatives aux 53 dernières masses râpées. Ces données sont rassemblées dans le tableau ci-après.

TABLEAU DES DONNÉES NUMÉRIQUES

R (origine à 21)	M (origine à 5)	T (origine à 67)	R	M	T
13,5	6,1	13,0	18,6	1,5	20,0
9,9	6,1	4,2	21,3	1,7	19,6
0,7	4,5	0,8	7,9	2,1	10,4
4,9	6,7	10,1	13,3	3,6	17,2
1,7	4,3	8,3	20,2	4,0	15,5
17,6	6,2	10,7	26,1	4,8	17,5
11,6	5,1	12,3	32,9	4,9	17,9
12,0	4,5	12,2	24,1	5,2	17,8
9,9	4,3	15,4	21,4	5,6	18,1
14,5	4,1	16,2	29,8	5,4	21,1
13,8	2,5	17,3	17,1	0,2	16,5
13,7	1,1	17,0	21,6	3,0	18,4
13,1	0,6	15,5	27,1	3,6	16,2
13,7	4,5	12,1	19,6	3,5	16,6
20,7	3,0	18,7	17,6	4,1	17,5
24,6	1,4	16,8	10,2	6,0	11,9
27,1	1,8	20,7	7,6	6,5	12,7
17,1	1,7	20,0	4,0	6,0	6,7
24,6	1,4	18,4	8,2	7,7	15,0
25,2	1,0	19,3	8,6	7,9	11,4
17,4	2,1	17,6	10,2	8,1	14,3
10,9	1,0	9,0	10,6	6,2	13,5
16,3	0,0	12,8	10,3	6,7	14,4
14,1	0,1	11,9	13,9	7,2	15,8
17,7	1,6	13,9	11,7	7,6	19,1
18,7	1,1	14,3	9,0	7,8	5,5
24,3	1,5	19,9			

Nous nous proposons de déterminer l'équation de régression linéaire :

$$R = \alpha M + \beta T + \text{Constante}$$

Nous nous attendons à trouver, compte tenu des remarques qui précèdent :

pour α une valeur négative ;

pour β une valeur positive.

D'après la théorie de la régression, α et β sont racines des équations :

$$\alpha \text{ var. } M + \beta \text{ cov. } MT = \text{cov. } R\, M ;$$

$$\alpha \text{ cov. } MT + \beta \text{ var. } T = \text{cov. } R\, T.$$

La résolution de ce système d'équations est simple mais nous allons utiliser une méthode qui se généralise facilement au cas où l'on étudie une régression par rapport à plus de deux variables et qui permet de calculer très aisément les variances des coefficients afin d'en tester la signification.

Cette méthode, qui fait appel à l'algèbre matricielle, permet de présenter et d'effectuer les calculs avec aisance et clarté.

Considérons le tableau carré symétrique :

$$\begin{Vmatrix} \text{var. } M & \text{cov. } M\,T \\ \text{cov. } M\,T & \text{var. } T \end{Vmatrix}$$

que nous appelons « matrice des variances et covariances ». Soit $\bar{\bar{a}}$ cette matrice.

$$\bar{\bar{a}} = \begin{Vmatrix} a_{11} & a_{12} \\ a_{12} & a_{22} \end{Vmatrix}$$

avec $a_{11} = \text{var. } M$

$a_{12} = \text{cov. } M\,T$

$a_{22} = \text{var. } T$

Désignons par $\bar{\bar{a}}^{-1}$ la matrice inverse de $\bar{\bar{a}}$

et posons : $\bar{\bar{a}}^{-1} = \begin{Vmatrix} a^{11} & a^{12} \\ a^{12} & a^{22} \end{Vmatrix}$

D'une manière générale, quand $\bar{\bar{a}}$ est une matrice carrée symétrique, le terme a^{ij} de la matrice inverse $\bar{\bar{a}}^{-1}$ est égal au quotient du mineur Aij du terme a_{ij} dans le déterminant $|a|$, par le déterminant $|a|$:

$$a^{ij} = \frac{Aij}{|a|} .$$

Dans le cas présent :

$$a^{11} = \frac{a_{22}}{a_{11}\, a_{22} - a_{12}^2} ,$$

$$a^{12} = \frac{-a_{12}}{a_{11}\, a_{22} - a_{12}^2} ,$$

$$a^{22} = \frac{a_{11}}{a_{11}\, a_{22} - a_{12}^2} ,$$

On a alors :

$$\alpha = a^{11} \text{ cov. } R\, M + a^{12} \text{ cov. } R\, T ,$$

$$\beta = a^{12} \text{ cov. } R\, M + a^{22} \text{ cov. } R\, T .$$

Dans l'exécution pratique des calculs, on considère la matrice des « sommes des carrés et produits centrés » dont les termes sont proportionnels aux estimations de var M, cov M T, var T, le coefficient de proportionnalité étant le nombre de degrés de liberté dont dépend la matrice.

Dans le cas présent, on a :

$$\Sigma R = 832,2 \qquad \bar{R} = 15,7018$$
$$\Sigma M = 209,2 \qquad \bar{M} = 3,9471$$
$$\Sigma T = 779,0 \qquad \bar{T} = 14,6981$$

$$\Sigma R^2 = 15.759,18 \qquad \Sigma (R - \bar{R})^2 = 2.692,07$$
$$\Sigma M^2 = 1.099,92 \qquad \Sigma (M - \bar{M})^2 = 274,17$$
$$\Sigma T^2 = 12.450,08 \qquad \Sigma (T - \bar{T})^2 = 1.000,25$$

$$\Sigma RM = 2.948,35 \qquad \Sigma (R - \bar{R}) \ (M - \bar{M}) = -336,48$$
$$\Sigma RT = 13.458,42 \qquad \Sigma (R - \bar{R}) \ (T - \bar{T}) = 1.226,65$$
$$\Sigma MT = 2.897,96 \qquad \Sigma (M - \bar{M}) \ (T - \bar{T}) = -176,89$$

D'où la matrice des sommes des carrés et produits centrés, matrice dépendant de 52 degrés de liberté.

Matrice des sommes des carrés et produits centrés
52 degrés de liberté

	M	T
M	274,17	— 176,89
T	— 176,89	1.000,25

Valeur du déterminant : 242.948,4704

Matrice inverse

10^{-3} 52 degrés de liberté

	M	T
M	4,117 128	0,728 097
T	0,728 097	1,128 511

On en déduit les coefficients α et β

$$\alpha = (10^{-3}\,4,117\ 128) \times (-336,48) + (10^{-3}\,0,728\ 097) \times (1.226,65)$$

$$\boxed{\begin{array}{l} \alpha = -\ 0,492\ 211 \\ \beta = +\ 1,139\ 297 \end{array}}$$

Ces résultats sont conformes à nos prévisions, quant au signe des coefficients.

Il en résulte que le rendement d'un moulin :

— diminuerait de presque un demi-kilogramme par heure lorsque l'excédent de mouillade augmente de 1 point ;

— augmenterait de plus de 1 kilogramme par heure lorsque la température augmente de 1 degré.

L'équation de régression s'écrit donc :

$$R - 15,7018 = -\ 0,492\ 211\ (M - 3,9471) + 1,139\ 297\ (T - 14,6981)$$

ou, avec les valeurs réelles :

$$R_0 - 36{,}7018 = - 0{,}492\ 211\ (M_0 - 8{,}9471) + 1{,}139\ 297\ (T_0 - 81{,}6981)$$

Test des coefficients α et β :

Décomposons la somme des carrés $\Sigma\,(R - \bar{R})^2 = 2.692{,}07$ en ses deux composantes orthogonales :

— somme des carrés due à la régression

— somme des carrés résiduels.

Comme $R - \bar{R} = \alpha(M - \bar{M}) + \beta(T - \bar{T}) + \varepsilon$,

ε désignant le résidu,

$$\Sigma\,(R - \bar{R})^2 = \alpha^2\Sigma\,(M - \bar{M})^2 + 2\,\alpha\beta\,\Sigma\,(M - \bar{M})(T - \bar{T}) - \beta^2\,\Sigma\,(T - \bar{T})^2 + \Sigma\,\varepsilon^2$$

ou

$$\Sigma\,(R - \bar{R})^2 = \alpha\left[\alpha\Sigma\,(M - \bar{M})^2 + \beta\Sigma\,(M - \bar{M})(T - \bar{T}\right]$$
$$+ \beta\left[\alpha\Sigma\,(M - \bar{M})(T - \bar{T}) + \beta\Sigma\,(T - \bar{T})^2\right]$$
$$+ \Sigma\,\varepsilon^2$$

D'après la définition même de α et β, le coefficient de α dans cette dernière expression est égal à $\Sigma\,(R - \bar{R})\,(M - \bar{M})$, celui de β à $\Sigma(R - \bar{R})\,(T - \bar{T})$.

La somme des carrées due à la régression est donc :

$$(0{,}492\ 211 \times 336{,}48) + (1{,}139\ 297 \times 1.226{,}65) = 1.563{,}14$$

D'où le tableau suivant d'analyse de la variance.

TABLEAU D'ANALYSE DE LA VARIANCE

Somme des carrés		Degrés de liberté	Carré moyen
Régression	1.563,14	2	
Résidu	1.128,93	50	22,5786
Total	2.692,07	52	

On obtient les variances des coefficients α et β en multipliant la variance résiduelle, soit 22,5786 par les termes de la diagonale principale de la matrice inverse.

$$\text{var}\ \alpha = (22{,}5786) \times (10^{-3}\ 4{,}117\ 128) = 0{,}092\ 959$$

$$\text{var}\ \beta = (22{,}5786) \times (10^{-3}\ 1{,}128\ 511) = 0{,}025\ 480$$

d'où les écarts-types σ_α et σ_β

$$\sigma_\alpha = 0{,}304$$

$$\sigma_\beta = 0{,}160$$

La valeur absolue de α vaut 1,62 fois son écart-type

— — β 7,12 — —

La valeur trouvée pour α ne peut donc être considérée comme significative. Par contre β est hautement significatif.

En conclusion, il apparaît que la fermentation a une influence très nette sur le rendement des moulins. L'excédent de mouillade ne semble pas, au contraire, avoir une action significative.

C'est pourquoi nous éliminons désormais la variable M, pour ne retenir que la variation de R en fonction de T

Détermination du rendement en fonction de la fermentation :

L'équation de régression s'écrit alors :

$$R - \bar{R} = \beta'(T - \bar{T})$$

$$\text{avec } \beta' = \frac{\Sigma(R - \bar{R})(T - \bar{T})}{\Sigma(T - \bar{T})^2} = \frac{1.226,65}{1.000,25} = 1,226\ 343$$

Avec les valeurs réelles, il vient :

$$\boxed{R_o = 1,226\ 343\ \ T_o - 63,4881}$$

C'est cette formule que nous utilisons désormais pour fixer le rendement du râpage.

Dans le cas présent, le tableau d'analyse de la variance s'établit comme suit :

TABLEAU D'ANALYSE DE LA VARIANCE

Somme des carrés		Degrés de liberté	Carré moyen
Régression	1.504,29	1	
Résidu	1.187,78	51	23,29
Total	2.692,07	52	

La variance de β' est égale à $\frac{23,29}{1000,25} = 0,023\ 284$ d'où $\sigma_{\beta'} = 0,152$ et $\frac{\beta'}{\sigma_{\beta'}} = 8,07$.

Il est intéressant de comparer les deux tableaux d'analyse de la variance :

Somme des carrés		Degrés de liberté	Carré moyen
Régression/T	1.504,29	1	
Régression/M,T	1.563,14	2	
Différence	58,85	1	58,85
Résidu	1.128,93	50	22,58

Le rapport des deux variances résiduelles $\frac{58,85}{22,58}$ est égal à 2,606.

La table de Snédécor — point 5 % — donne pour la colonne 1 et la ligne ∞ le rapport 3,84. La valeur observée étant inférieure à ce rapport, on est en droit de conclure que ces deux variances ne sont pas significativement distinctes. Autrement dit, la régression par rapport à M et T n'apporte pas dans la détermination de R une précision meilleure que celle donnée par la régression par rapport à T seul. Ceci confirme les conclusions précédentes.

Régression curviligne. — Polynômes orthogonaux.

Sur le graphique ci-joint (Fig. 1), nous avons porté le nuage de points (les rendements en ordonnée, les températures en abscisses) et avons tracé la droite de régression d'équation

$$R_0 = 1,226\ 343\ T_0 - 63,4881$$

Il apparaît que le nuage de points a une forme légèrement concave, la concavité étant vers le coin supérieur gauche du graphique.

C'est pourquoi, nous avons tenté d'améliorer la détermination de R en fonction de T en prenant, au lieu d'une formule linéaire, une fonction du 2^e, ou même du 3^e degré en T.

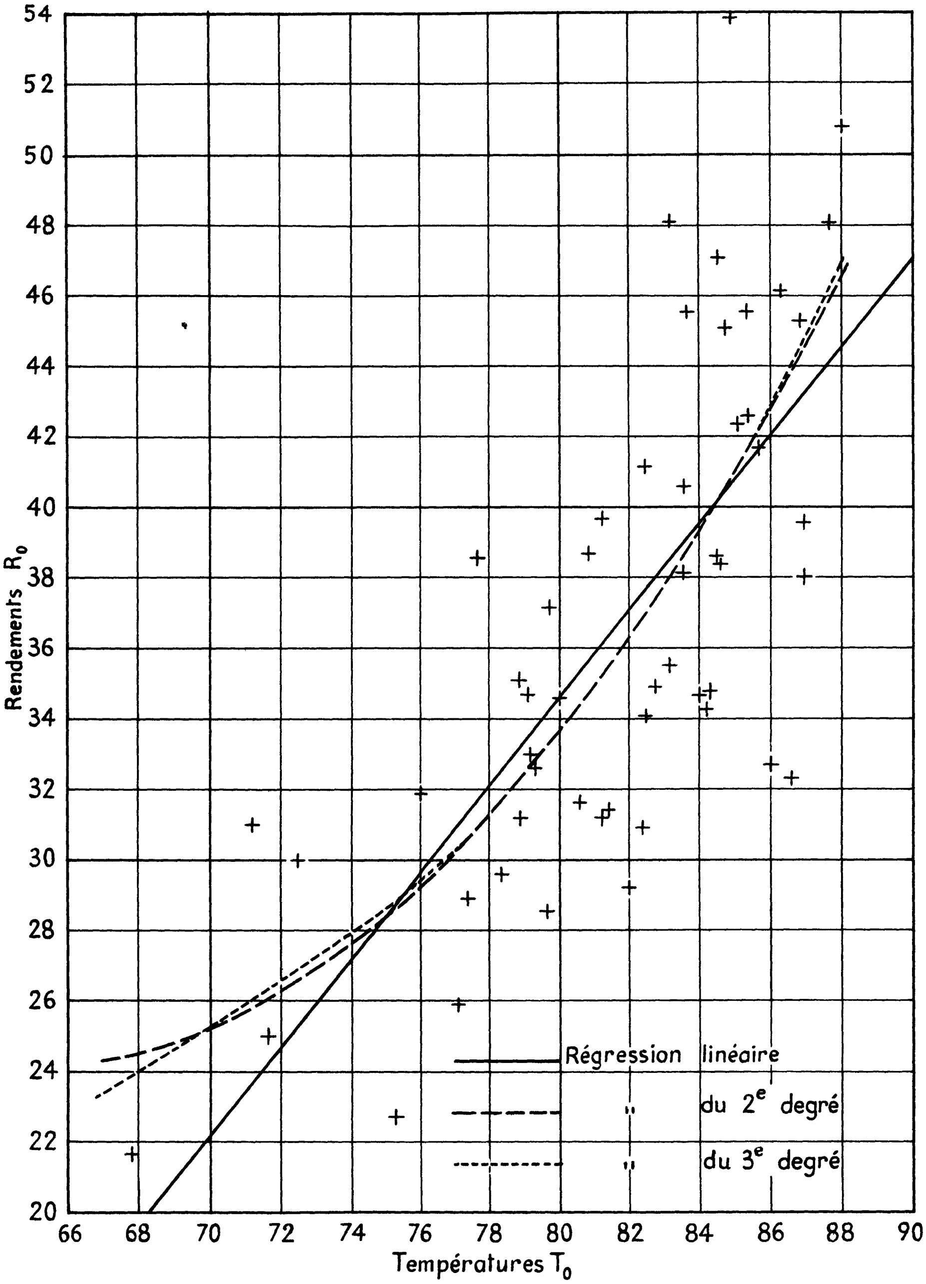

Fig. 1

NOTA : Les courbes du 2^e et du 3^e degré sont pratiquement confondues entre 78° et 85°

Nous avons adopté la méthode des polynômes orthogonaux de Gauss, dont nous rappelons brièvement le principe.

Nous écrivons l'équation de la ligne de régression sous la forme :

$$R = b_0 P_0 + b_1 P_1 + \dots + b_p P_p,$$

où les P sont des polynômes en T, P_j étant du degré j.

On détermine les polynômes P de manière que :

$$\Sigma(P_j\ P_k) = o \quad (j \neq k)$$

la sommation s'étendant aux valeurs observées.

En minimisant

$\Sigma(R - b_0 P_0 - b_1 P_1 \ldots - b_p P_p)^2$ par rapport aux coefficients b nous obtenons des équations telles que :

$$\Sigma(R\ P_j) - b_0\ \Sigma(P_0\ P_j) - \ldots - b_p\ \Sigma(P_p\ P_j) = o$$

qui, en vertu de la relation d'orthogonalité, se réduisent à :

$$\Sigma(R\ P_j) - b_j\ \Sigma(P_j^2) = o$$

Cette méthode présente les deux avantages suivants :

a) lorsqu'on possède l'équation de régression de degré p, on obtient celle de degré p + 1 en ajoutant $b_{p+1}\ P_{p+1}$;

b) la somme des carrés résiduels est :

$$\begin{aligned} U &= \Sigma(R - b_0\ P_0 - \ldots - b_p\ P_p)^2 \\ &= \Sigma R^2 - b_0^2\ \Sigma(P_0^2) - \ldots - b_p^2\ \Sigma(P_p^2). \end{aligned}$$

L'effet d'un terme $b_j\ P_j$ est de réduite U de la quantité $b_j^2\ \Sigma\ (P_j^2)$ et on peut examiner si ce terme apporte une diminution de U significative.

L'exemple que nous traitons permet de voir comment peuvent être calculés les polynômes P et les coefficients b.

Nous nous proposons de calculer l'équation de la ligne de régression du 3ᵉ degré :

$$R = b_0 P_0 + b_1 P_1 + b_2 P_2 + b_3 P_3$$

Nous aurons besoin des quantités suivantes :

ΣT	=	779,0	Σn	=	53
ΣT^2	=	12.450,08	ΣR	=	832,2
ΣT^3	=	208.059,704	ΣR^2	=	15.759,18
ΣT^4	=	3.588.029,6300	ΣRT	=	13.458,42
ΣT^5	=	63.381.042,31040	ΣRT^2	=	227.742,658
ΣT^6	=	1.141.240.047,044828	ΣRT^3	=	3.970.768,8312

Les expressions des polynômes sont les suivantes :

$$P_0 = 1$$

$$P_1 = \frac{\begin{vmatrix} n & \Sigma T \\ 1 & T \end{vmatrix}}{n}$$

$$P_2 = \frac{\begin{vmatrix} n & \Sigma T & \Sigma T^2 \\ \Sigma T & \Sigma T^2 & \Sigma T^3 \\ 1 & T & T^2 \end{vmatrix}}{\begin{vmatrix} n & \Sigma T \\ \Sigma T & \Sigma T^2 \end{vmatrix}}$$

$$P_3 = \frac{\begin{vmatrix} n & \Sigma T & \Sigma T^2 & \Sigma T^3 \\ \Sigma T & \Sigma T^2 & \Sigma T^3 & \Sigma T^4 \\ \Sigma T^2 & \Sigma T^3 & \Sigma T^4 & \Sigma T^5 \\ 1 & T & T^2 & T^3 \end{vmatrix}}{\begin{vmatrix} n & \Sigma T & \Sigma T^2 \\ \Sigma T & \Sigma T^2 & \Sigma T^3 \\ \Sigma T^2 & \Sigma T^3 & \Sigma T^4 \end{vmatrix}}$$

En remplaçant dans ces expressions : n, ΣT, ΣT^2, etc... par les valeurs indiquées plus haut, on obtient :

$$P_0 = 1$$

$$P_1 = T - 14{,}698\ 113$$

$$P_2 = T^2 - 25{,}060\ 754\ T + 133{,}438\ 633$$

$$P_3 = T^3 - 34{,}797\ 896\ T^2 + 342{,}244\ 77\ T - 781{,}\ 779\ 03$$

Les coefficients b sont obtenus par le quotient de deux déterminants.

D'une manière générale :

$$b_p = \frac{\Delta_p^{(p)}}{\Delta^{(p)}}$$

$\Delta^{(p)}$ et $\Delta_p^{(p)}$ ayant les expressions suivantes :

$$\Delta^{(p)} = \begin{vmatrix} n & \Sigma T & \Sigma T^2 \ldots\ldots\ldots\ldots & \Sigma T^p \\ \Sigma T & \Sigma T^2 & \Sigma T^3 \ldots\ \ldots\ldots\ldots & \Sigma T^{p+1} \\ \vdots & & & \\ \Sigma T^p & \Sigma T^{p+1} & \Sigma T^{p+2} & \Sigma T^{2p} \end{vmatrix}$$

$\Delta_p^{(p)}$ s'obtient en remplaçant, dans l'expression $\Delta^{(p)}$ les termes de la dernière colonne par :

$$\begin{matrix} \Sigma R \\ \Sigma RT^2 \\ \Sigma RT \\ \vdots \\ \Sigma RT^p \end{matrix}$$

Ainsi, dans le cas présent, l'expression de b_3 est la suivante :

$$b_3 = \frac{\begin{vmatrix} 53 & 779{,}0 & 12.450{,}08 & 832{,}2 \\ 779{,}0 & 12.450{,}08 & 208.059{,}704 & 13.458{,}42 \\ 12.450{,}08 & 208.059{,}704 & 3.588.029{,}6300 & 227.742{,}658 \\ 208.059{,}704 & 3.588.029{,}6300 & 63.381.042{,}31040 & 3.970.768{,}8312 \end{vmatrix}}{\begin{vmatrix} 53 & 779{,}0 & 12.450{,}08 & 208.059{,}704 \\ 779{,}0 & 12.450{,}08 & 208.059{,}704 & 3.588.029{,}6300 \\ 12.450{,}08 & 208.059{,}704 & 3.588.029{,}6300 & 63.381.042{,}31040 \\ 208.059{,}704 & 3.588.029{,}6300 & 63.381.042{,}31040 & 1.141.240.047{,}044828 \end{vmatrix}}$$

Les calculs se font en développant les déterminants par rapport à leur dernière colonne ; les mineurs ont déjà été tous calculés en établissant les polynômes P.

Nous obtenons :

$$b_0 = 15{,}701\ 886$$
$$b_1 = 1{,}226\ 344$$
$$b_2 = 0{,}042\ 936$$
$$b_3 = 0{,}001\ 230\ 6$$

En définitive, les équations de régression sont les suivantes :

Régression linéaire :

$$R = 1{,}226344\ T - 2{,}323\ 057$$

nous retrouvons — bien entendu — l'équation de régression établie précédemment.

Régression du 2e degré :

$$R = 0{,}042\ 936\ T^2 + 0{,}150\ 335\ T + 3{,}406\ 264$$

Régression du 3e degré :

$$R = 0{,}001\ 231\ T^3 + 0{,}000\ 114\ T^2 + 0{,}571\ 495\ T + 2{,}444\ 222$$

Les courbes de régression du 2e et du 3e degré ont également été tracées sur le graphique. Ces courbes sont pratiquement confondues entre elles. Leur concavité est bien orientée comme nous l'avions envisagé.

Enfin, elles ne s'éloignent pas beaucoup de la droite.

La somme des carrés résiduels, que nous avons appelée U, diminue de

$$b_p^2\ \Sigma\,(P_p^2)$$

lorsque l'on passe de la courbe de régression de degré p — 1 à celle de degré p. On établit que :

$$b_p^2\ \Sigma\,(P_p^2) = b_p^2\ \frac{\Delta^{(p)}}{\Delta^{(p-1)}}$$

La somme totale des carrés est $\Sigma R^2 = 15.759{,}18$.

$b_0\ P_0$ L'effet de ce terme est de diminuer ΣR^2 de la quantité $b_0^2\ \Delta^{(0)}$ ou

$$(15{,}701\ 886)^2 \times 53 = 13.067{,}101\ 529$$

et

$$U_0 = 15.759{,}18 - 13.067{,}101\ 529 = 2.692{,}070\ 471$$

On retrouve l'expression de $\Sigma(R - \overline{R})^2$, puisque $b_0\ P_0 = \overline{R}$.

$b_1\ P_1$ Avec le terme de 1er degré, U_0 diminue de :

$$b_1^2\ \frac{\Delta^{(1)}}{\Delta^{(0)}} = (1{,}226\ 344)^2 \times 53.013{,}24 \times \frac{1}{53} = 1.504{,}2952$$

et $U_1 = 1.187{,}7753$

$b_2\ P_2$ On trouve, de la même manière, que le terme du 2e degré a pour effet de diminuer U de :

$$64{,}9247$$

et $U_2 = 1.122{,}8506$

$b_3\ P_3$ Le polynôme du 3e degré ne diminue U que de :

$$1{,}5760$$

et $U_3 = 1.121{,}2746$

Ces résultats sont rassemblés dans le tableau d'analyse de la variance ci-après :

Somme des carrés		Somme des carrés résiduels	Degrés de liberté	Variance résiduelle
$\Sigma(R - \bar{R})^2$	2.692,0705		52	
Régression linéaire. . .	1.504,2952	1.187,7753	51	23,29
Polynôme du 2ᵉ degré	64,9247	1.122,8506	50	22,46
Polynôme du 3ᵉ degré .	1,5760	1.121,2746	49	22,88

Pour tester l'influence du polynôme du 2ᵉ degré, nous devons comparer 64,9247 à la variance résiduelle 22,46. Le rapport de ces deux variances est 2,89, valeur inférieure au rapport 3,84 indiqué par la table de Snédécor à l'intersection de la colonne 1 et de la ligne ∞ (pour P = 0,05). Donc, le fait d'ajouter le polynôme $b_2 P_2$ ne diminue pas de façon significative la somme des carrés résiduels. L'adoption d'une formule du 2ᵉ degré en T n'apporte aucune précision supplémentaire, et on peut se limiter à la formule de régression linéaire.

Quant au polynôme du 3ᵉ degré, il laisse pratiquement inchangée la somme des carrés résiduels. C'est ce qui se traduit sur le graphique par le fait que les 2 courbes sont presque confondues.

Conclusion :

Cette étude a mis en évidence et surtout a permis de mesurer l'influence de la température atteinte par le tabac au cours de sa fermentation en masses sur le rendement des moulins du râpage.

Nous adoptons désormais, pour fixer ce rendement, une formule linéaire en T.

L'adjonction du paramètre M, ou l'adoption d'une formule du 2ᵉ ou du 3ᵉ degré en T n'apporte aucune amélioration sensible dans la détermination de R.

Actuellement, cette formule est appliquée depuis près d'un an et elle donne satisfaction. Le rendement réalisé est voisin du rendement exigé et souvent même légèrement supérieur.

En définitive, cette étude statistique a permis de rémunérer les ouvriers râpeurs d'une manière plus équitable ; les mesures prises ont stimulé les ouvriers désormais intéressés par la prime de rendement, qui, autrefois, dépendait de facteurs indépendants de leur activité.

ANNEXE

TRACÉ DES COURBES DE RÉGRESSION

1° Régression linéaire :

$T_0 = 81{,}7$ $\qquad R_0 = 36{,}7$

$T_0 = 91{,}7$ $\qquad R_0 = 49{,}0$

$T_0 = 71{,}7$ $\qquad R_0 = 24{,}4$

2° Régression du 2ᵉ degré :

$T_0 = 67$	$R_0 = 24{,}4$	$T_0 = 81$	$R_0 = 34{,}9$
69	24,9	82	36,3
72	26,2	83	37,8
74	27,6	84	39,4
75	28,4	85	41,0
77	30,2	86	42,8
78	31,3	87	44,6
79	32,4	88	46,5
80	33,6		

3° Régression du 3ᵉ degré :

$T_0 = 67$	$R_0 = 23{,}4$	$T_0 = 80$	$R_0 = 33{,}6$
69	24,6	82	36,2
71	25,8	84	39,2
72	26,5	87	44,8
75	28,7	88	46,9
77	30,4		