J. ULMO

F. BASTENAIRE

Analyse d'un plan factoriel d'expérimentation dont les résultats sont qualitatifs

Revue de statistique appliquée, tome 5, nº 2 (1957), p. 13-18

<http://www.numdam.org/item?id=RSA_1957__5_2_13_0>

ANALYSE
D'UN PLAN FACTORIEL D'EXPÉRIMENTATION DONT LES RÉSULTATS SONT QUALITATIFS

par

J. ULMO et F. BASTENAIRE

Statisticiens à l'Institut de Recherches de la Sidérurgie

Il arrive que dans un premier dépouillement rapide des résultats d'une expérience planifiée de façon factorielle, pour examiner l'influence d'un facteur contrôlé, on soit conduit à procéder de la façon suivante :

a) Classer les résultats en deux catégories I ou II;

b) Calculer les proportions de résultats de l'une des catégories I par exemple, observés pour chacun des niveaux $A_1, A_2, \ldots, A_r$ *du facteur contrôlé;*

c) Comparer ces proportions observées $X_1, X_2, \ldots, X_r$.

La comparaison de ces proportions observées ne parait pas à priori relever des méthodes classiques de comparaison de proportions car chacune d'elles est calculée sur un ensemble N de résultats qui ne proviennent pas d'une seule population mais de k populations, à priori différentes puisqu'elles correspondent chacune à des combinaisons différentes des autres facteurs contrôlés B, C, D.. à raison de ν_i *résultats par combinaison.*

Ce que l'on sait toutefois, c'est qu'aux différents niveaux $A_1, A_2, \ldots, A_r$ *du facteur A correspondent les mêmes combinaisons des autres facteurs contrôlés et les mêmes nombres* ν_i $(i = 1, 2, \ldots, k)$ *de résultats par combinaison.*

L'hypothèse que l'on désire tester est donc que les probabilités $p_1, p_2, \ldots, p_k$ *des résultats de la catégorie I correspondant aux k combinaisons des autres facteurs contrôlés ne dépendent pas du niveau du facteur A.*

L'étude qui suit montre que les méthodes classiques de comparaison de proportions tout en n'étant pas les plus efficaces, c'est-à-dire en correspondant à un risque de première espèce inférieur à celui pour lesquelles elles sont établies, sont toutefois valables.

I - ÉTUDE DE LA DISTRIBUTION DE X

Soit n_i $_{i=1,2,\ldots k}$ le nombre de résultats de la catégorie considérée I, obtenus par tirage de ν_i résultats dans l'urne U_i de composition p_i , $q_i = 1 - p_i$.

n_i a une distribution binomiale de moyenne $\nu_i\, p_i$, variance $\nu_i\, p_i\, q_i$.

La proportion des résultats de la catégorie I obtenus en faisant ν_i tirages dans chacune des k urnes est :

$$X = \frac{\sum_{i=1}^{k} n_i}{N} = \frac{S}{N} \quad \text{où } S = \sum_{i=1}^{k} n_i \text{ est le nombre total}$$

des résultats de la catégorie I et $N = \sum_{i=1}^{k} \nu_i$ le nombre total des tirages.

On a
$$E(S) = \sum_i \nu_i\, p_i$$

et
$$\sigma_S^2 = \sum_i \nu_i\, p_i\, q_i$$

et, par suite, $E(X) = \dfrac{\sum_i \nu_i\, p_i}{N} = p.$ en posant $p. = \dfrac{\sum_i \nu_i\, p_i}{\sum_i \nu_i}$

$p.$ est la proportion moyenne de la catégorie I dans l'ensemble constitué par la réunion de ν_1 urnes U_1 , ν_2 urnes U_2 ,.... ν_k urnes U_k

Si $\nu_1 = \nu_2 , \ldots = \nu_k = \nu$:

$p. = \frac{\sum_i p_i}{k}$ est la proportion moyenne de la catégorie I dans l'ensemble des k urnes.

et $$\sigma_X^2 = \frac{\sigma_S^2}{N^2} = \frac{\sum_i \nu_i \; p_i \, q_i}{N^2} = \frac{1}{N} \; \frac{\sum_i \nu_i \; p_i \; q_i}{\sum_i \nu_i}$$

n_i étant déjà la somme de ν_i valeurs indépendantes d'une variable aléatoire x_i prenant la valeur 1 si on tire un résultat de la catégorie I et la valeur 0 si on tire un résultat de la catégorie II, S est la somme de N telles valeurs et X leur valeur moyenne. Par suite, si N n'est pas trop faible et si l'ensemble des p_i n'est pas trop voisin de 0 ou 1 : les distributions de S et de X sont très sensiblement normales.

II - CALCUL D'UNE LIMITE SUPÉRIEURE DE L'ÉCART-TYPE DE X QUI NE DÉPENDE QUE DE P

Si l'on connaissait σ_X , le problème de la comparaison de r valeurs observées X_1 , X_2 , ..., X_r serait très simple puisqu'il se ramènerait au test de l'hypothèse d'une même moyenne pour r variables normales indépendantes, d'écarts-types connus, test qui consiste à regarder si

$$\sum_{i=1}^{r} \frac{(X - \overline{X})^2}{\sigma_{X_i}^2} \qquad \text{avec} \qquad \overline{X} = \frac{\sum_i \frac{X_i}{\sigma_{X_i}^2}}{\sum_i \frac{1}{\sigma_{X_i}^2}}$$

ne diffère pas significativement de X^2 avec r-1 degrés de liberté. (Dans le cas où r = 2, il revient au même de regarder si $\frac{X_1 - X_2}{\sqrt{\sigma_{X_1}^2 + \sigma_{X_2}^2}}$ ne diffère pas significativement d'une variable normale réduite).

Mais étant donné l'expression de σ_X , la connaissance de la seule valeur X ne nous permet pas d'estimer σ_X . Il faudrait pouvoir estimer les p_i ce que, précisément, nous ne voulons pas faire et ne pourrions du reste pas faire de façon valable si les ν_i sont petits comme c'est généralement le cas (par exemple $\nu_i = 1$).

Nous avons alors cherché à trouver une limite supérieure σ_{XL} de σ_X dont la valeur puisse être estimée, donc, ne dépendant que de p. que l'on peut estimer par X $[E(X) = p.]$. On verra dans III/a que la précision de cette estimation est au moins égale à celle du paramètre d'une distribution binomiale à partir de la fréquence relative obtenue sur N tirages.

On voit, en effet, que si l'hypothèse nulle est rejetée avec σ_{XL} , elle le sera a fortiori avec σ_X . Par contre, l'hypothèse nulle peut être acceptée avec σ_{XL} sans l'être avec σ_X . En d'autres termes, en remplaçant σ_X par σ_{XL} , on obtient un test dont le risque de première espèce est inférieur au risque "nominal" pour lequel il a été établi.

Une première limite supérieure de σ_X^2 vient immédiatement à l'esprit. C'est celle que l'on obtient en majorant $p_i \, q_i = p_i \, (1 - p_i)$ par 1/4 soit 1/4 N.

Mais on peut aussi tenir compte du fait que l'on connait la valeur de $p_. = \frac{\sum \nu_i \; p_i}{N}$ et chercher la limite supérieure des valeurs de σ^2_X correspondant à une valeur donnée p. de $\frac{\sum \nu_i \; p_i}{N}$.

On peut écrire :

$$\sigma^2_X = \frac{1}{N^2} \sum_i \nu_i \; p_i \; q_i = \frac{1}{N^2} \sum \nu_i \; p_i \; (1 - p_i)$$

$$= \frac{1}{N^2} \left[N \, p_. - \sum \nu_i \; p_i^2 \right]$$

On voit donc que chercher le maximum de σ^2_X revient à chercher le minimum de $\sum \nu_i \; p_i^2$ compte tenu du fait que $\sum \nu_i \; p_i = N \, p_. =$ Cte.

La méthode des multiplicateurs de Lagrange qui consiste à déterminer les valeurs des p_i et de λ rendant minimum

$\Omega^2 = \sum \nu_i \; p_i^2 - \lambda \sum \nu_i \; (p_i - p_.)$ donne $\frac{\partial \Omega^2}{\partial p_i} = 2 \, \nu_i \; p_i - \lambda \, \nu_i$,

soit $p_i = \lambda / 2$ et, par suite $p_1 = p_2, \ldots\ldots, = p_k = p_.$.

On a donc :

$$\sigma^2_X \leqslant \sigma^2_{X_L} = \frac{p_. \; (1 - p_.)}{N} = \frac{p_. \; q_.}{N}.$$

On remarquera que σ^2_{XL} n'est autre que la variance de la fréquence relative quand on fait N tirages dans une urne à 2 catégories de composition p. q. tandis que E (X) = p. est la valeur moyenne de cette même fréquence relative.

<u>X a donc même moyenne et est moins dispersée que la valeur que l' on obtiendrait en tirant au hasard dans une urne unique de composition</u> p. q.

III - PROBLÈMES DE JUGEMENT SUR ÉCHANTILLONS RELATIFS A X

A) DETERMINATION D'UN INTERVALLE DE CONFIANCE A UN RISQUE α DONNÉ POUR p. .

L'intervalle cherché sera certainement inférieur à l'intervalle obtenu en remplaçant σ_X par σ_{X_L} c'est-à-dire, en estimant σ_{X_L} par $\sqrt{\frac{X\,(1 - X)}{N}}$, à l'intervalle $X \pm t_{\alpha/2} \sqrt{\frac{X\,(1-X)}{N}}$ où $t_{\alpha/2}$ désigne la limite de l'intervalle de probabilité au même risque α pour une variable normale réduite.

B) COMPARAISON DE DEUX PROPORTIONS OBSERVEES X_1 ET X_2

L'hypothèse à tester est, comme on l'a vu au début, celle de l'identité des probabilités p_{i_1} et p_{i_2} (i = 1,2,....,k) des urnes U_{i_1} et U_{i_2} correspondant respectivement aux deux proportions observées.

Dans ces conditions, on a, non seulement $E\,(X_1) = E\,(X_2)$ mais aussi $\sigma_{X_1} = \sigma_{X_2} = \sigma_X$ et, par suite $\frac{X_1 - X_2}{\sigma_X \sqrt{2}}$ doit être distribué comme une variable normale réduite.

En remplaçant σ_X par σ_{X_L} que l'on estimera par $s_{X_L} = \sqrt{\frac{\overline{X}\,(1 - \overline{X})}{N}}$ avec $\overline{X} = \frac{\overline{X}_1 + \overline{X}_2}{2}$ meilleure estimation de $p_.$; on obtient un test qui correspond à un risque de première espèce au plus égal au risque nominal pour lequel il a été calculé.

On remarquera que ce test n'est autre que le test classique de comparaison de deux proportions que l'on peut encore écrire sous la forme :

$$(1) \qquad \frac{N\ (X_1 - X_2)^2}{2\overline{X}\ (1 - \overline{X})}$$

doit être distribué comme χ^2 avec 1 degré de liberté.

Ceci n'a rien d'étonnant puisque $\overline{X}$ est l'estimation du χ^2 minimum de p. avec

$$(2) \qquad \chi^2 = \frac{(X_1 - p.)^2 + (X_2 - p.)^2}{\frac{p.\ (1 - p.)}{N}}$$

Ceci montre que si la quantité (2) est distribuée comme χ^2 avec 2 degrés de liberté, celle que l'on obtient en remplaçant p. par $\overline{X}$, soit (1) , est distribuée comme χ^2 avec 1 degré de liberté.

C) COMPARAISON DE r PROPORTIONS OBSERVEES X_1 , X_2 ......, X_r

L'hypothèse à tester est celle de l'identité des probabilités p_{i_1} , p_{i_2} , ...p_{i_r} (i = 1,2 , . . . , k) des urnes U_{i_1} , U_{i_2} . . . U_{i_r} correspondant respectivement aux r proportions observées.

On a donc encore, non seulement $E\ (X_1) = E\ (X_2), \ldots = E\ (X_r)$ mais aussi $\sigma_{X_1} = \sigma_{X_2} = \sigma_{X_r} = \sigma_X$ et, par suite :

$$\frac{\Sigma\ (X_i - \overline{X})^2}{\sigma_X^2} \text{ avec } \overline{X} = \frac{X_1 + X_2 , \ldots\ldots , X_r}{r}$$

doit être distribué comme χ^2 avec r-1 degrés de liberté.

En remplaçant σ_X^2 par $\sigma_{X_L}^2$ que l'on estimera par $\frac{\overline{X}\ (1 - \overline{X})}{N}$, on obtient un test qui correspond à un risque de première espèce au plus égal au risque nominal pour lequel il a été calculé et qui n'est autre que le test classique de comparaison de r proportions comme on le voit quand on l'écrit sous la forme :

$$\frac{N\ \Sigma\ (X_i - \overline{X})^2}{\overline{X}\ (1 - \overline{X})} = \frac{N}{\overline{X}\ (1 - \overline{X})} \left[\Sigma\ X_i^2 - N\ \overline{X}^2 \right] = \frac{\sum_i (S_i - S_i')^2}{N\ p'\ q'}$$

doit être distribué comme χ^2 avec r-1 degrés de liberté.

$p' = \overline{X}$ est l'èstimation de p.

$q' = 1 - p'$

et $S_i' = N\ \overline{X} = N\ p'$ est l'estimation de $S_i = N\ X_i$ fréquence des résultats de la catégorie I.

Le fait que l'on retombe sur le test classique résulte de ce que X est l'estimation du χ^2 minimum de p. avec

$$\chi^2 = \frac{N\ \Sigma\ (X_i - p.)^2}{p.\ (1 - p.)}$$

ce qui justifie l'obtention d'une quantité distribuée comme χ^2 avec r-1 degrés de liberté en remplaçant p. par $\overline{X}$

EXEMPLE :

Etude de l'influence d'un certain nombre de facteurs sur l'adhérence de la calamine de feuillards d'acier.

L'adhérence de la calamine a été notée de 0 à 5 d'après l'observation de coupes :

0 désignant une calamine peu adhérente, et
5 une calamine très adhérente.

Les différents facteurs à étudier ont été combinés factoriellement, et il n'y a pas eu de répétition.

Ce sont :

la qualité de l'acier, 3 niveaux :	(T (O (M
la température de laminage, 2 niveaux :	(1 la plus élevée (2
la position du prélèvement sur les couronnes de feuillards, 3 niveaux :	(tête T (milieu M (pied P
l'usine ayant procédé au laminage final, 3 usines :	(I (II (III

Il n'a en effet, été utilisé qu'une seule coulée de chaque qualité d'acier, les trois coulées ayant été élaborées et laminées à la suite en demi produits(brames) par une même usine afin d'assurer la plus grande homogénéité possible au matériel de base.

Afin d'obtenir un classement des résultats en deux catégories seulement, ceux-ci ont été regroupés de la façon suivante :

Nous avons noté + les résultats de note 4 ou 5
et - les résultats de note 0, 1, 2 ou 3
de façon à avoir des proportions marginales de + qui ne soient pas trop voisines de 0 ou 1 pour que l'approximation normale puisse être acceptable malgré le petit nombre des essais.

Les résultats obtenus pour les 54 combinaisons factorielles sont les suivants :

Qualité de l'acier

		T			O			M			Proportion de + : x
	Usine	I	II	III	I	II	III	I	II	III	
température 1	T	+	+	-	+	-	-	-	-	+	$\frac{12}{27}$
	M	+	+	-	+	+	-	-	-	-	
	P	+	+	-	+	-	-	-	+	-	
température 2	T	+	+	+	+	+	+	+	+	-	$\frac{21}{27}$
	M	-	+	+	+	+	-	+	+	-	
	P	+	+	+	+	+	-	+	+	-	
Proportion de + : x		$\frac{14}{18}$			$\frac{11}{18}$			$\frac{8}{18}$			

Proportions de + :

Usine I	$\frac{14}{18}$	Position du prélèvement :	T	$\frac{12}{18}$
Usine II	$\frac{14}{18}$		M	$\frac{10}{18}$
Usine III	$\frac{5}{18}$		P	$\frac{11}{18}$

Les résultats des comparaisons des proportions marginales obtenues pour les différents niveaux de chacun des facteurs sont rassemblés dans le tableau suivant :

Facteur	Moyenne $\overline{x}$	$\dfrac{N\Sigma(X_i - \overline{x})^2}{\overline{x}\ (1 - \overline{x})}$	degrés de liberté	Signification (1)
température de laminage	$\frac{11}{18}$	6,31	1	S
Usine	$\frac{11}{18}$	23,1	2	THS
Qualité d'acier	$\frac{11}{18}$	4,2	2	nS
Position de prélèvement	$\frac{11}{18}$	0,46	2	nS

CONCLUSIONS

1) L'usine, c'est-à-dire son train de laminoirs a une influence prépondérante. (Les usines I et II dont les trains sont peu différents, ne donnent pas des résultats significativement différents, tandis que l'usine III donne une calamine nettement plus adhérente).

2) La température de laminage a également une influence significative, l'adhérence diminuant quand celle-ci augmente.

(1) S = significatif au risque 5 %
THS = significatif au risque de 0,1 %
nS = non significatif au risque 5 %