

APPLICATION DU TEST DE COMPARAISON DE DEUX MOYENNES D'UNE SÉRIE ORDONNÉE AU CLASSEMENT DES TAUREAUX SUR DESCENDANCE

A. ARNAUD

*Station centrale de Génétique animale,
Centre national de Recherches zootechniques, Jouy-en-Josas (Seine-et-Oise)*

SOMMAIRE

Le test de comparaison de moyennes ordonnées (DUNCAN) est appliqué à la comparaison des performances moyennes des descendants de taureaux d'insémination pris 2 à 2.

Une application *a posteriori* est effectuée sur les résultats relatifs à une série de taureaux. Une recherche *a priori* de l'effectif de descendants à contrôler pour faire apparaître une différence significative aux seuils de 10 p. 100, 5 p. 100 et 1 p. 100 entre 2 taureaux de rangs voisins a été réalisée dans une deuxième partie.

I. — BUT DE L'ÉTUDE

Les travaux de génétique biométrique appliqués au progeny test des taureaux d'insémination, ont été consacrés essentiellement ;

— à l'établissement d'index génotypiques de taureaux d'après les performances d'un certain nombre de descendants ;

— à la recherche d'un plan de sélection par testage, correspondant au progrès génétique ou économique maximum par unité de temps.

Ces travaux présentent un aspect synthétique, en ce sens qu'il se rapportent à l'ensemble de la population. Ils concernent, d'autre part, des résultats moyens et ne se réfèrent pas à la probabilité associée à ces résultats. En pratique, les décisions à prendre pour le choix des taureaux, nécessitent souvent des comparaisons analytiques des descendances des taureaux pris 2 à 2 : comparaisons dont on souhaite par ailleurs connaître la signification statistique.

Le but de notre étude est précisément :

- d'interpréter statistiquement les résultats de progeny-test par une comparaison analytique des descendances de taureaux ;
- de chercher, *a priori*, les valeurs de l'effectif des lots de testage, permettant de réaliser un classement précis des géniteurs.

II. — MÉTHODE D'ÉTUDE

Dans la pratique, les résultats de testage se présentent sous forme d'une série de moyennes ordonnées, le nombre de descendants par taureaux étant variable pour chaque échantillon.

Dans ce cas, la différence entre deux moyennes m_i et m_j ne suit plus la loi de probabilité de t , qui est basée sur la comparaison de deux échantillons indépendants, et tirés au hasard, mais une loi de probabilité qui dépend entre autres du nombre de moyennes que le rangement par ordre de valeur croissante a intercalé entre les deux moyennes.

Parmi différents tests proposés, le test de DUNCAN (2) semble être bien adapté à la résolution de ce problème (RIVES M., 1959). En outre, KRAMER C. Y., (1956), a donné une extension du test de DUNCAN pour le cas où le nombre de répétitions est variable pour chaque échantillon.

Dans notre étude, nous utiliserons donc la méthode proposée par DUNCAN et étendue par KRAMER C. Y. au cas d'effectifs variables par groupes :

Soit une série de moyennes ordonnées :

$$\bar{x}_a ; \bar{x}_b \dots \bar{x}_i \dots \bar{x}_k \dots ; \bar{x}_n.$$

Les deux moyennes \bar{x}_i et \bar{x}_k seront significativement différentes si nous avons :

$$(\bar{x}_i - \bar{x}_k) \sqrt{\frac{2n_i \cdot n_k}{n_i + n_k}} > R_p$$

avec $R_p = S \cdot Z_{p,n}$

avec

\bar{x}_i = moyenne des descendants du i ème taureau.

n_i = nombre de descendants contrôlés du i ème taureau.

S^2 = variance due à l'erreur.

p = nombre d'intervalles plus 1 compris entre les deux moyennes.

n = nombre de degrés de liberté de l'erreur.

$Z_{p,n}$ = étendue, après transformation de STUDENT, significative pour un nombre d'intervalles p et n degrés de liberté.

(Voir table de DUNCAN publiée par L. H. HARTER, 1960).

Remarque :

Dans le cas où nous comparons deux moyennes seulement, le test t et le test de DUNCAN donnent des résultats identiques (KRAMER C. Y., 1956).

III. — INTERPRÉTATION DES RÉSULTATS DE TESTAGE

L'application *a posteriori* du test de DUNCAN à une série de onze taureaux du centre d'insémination artificielle de Marmillat, comparés sur le poids à 90 jours de leurs descendants, a fourni les résultats suivants (tabl. 1) (Résultats corrigés pour l'influence du sexe).

Variance due à l'erreur $S^2 = 419,04$.

Degrés de libertés de l'erreur, $n = 256$.

En nous reportant à la table de DUNCAN, nous calculons, pour chaque valeur de p et pour les seuils de probabilité utilisés : (0,10 ; 0,05 ; 0,01) les valeurs correspondantes de $R_p = S \cdot Z_p$, n (tabl. 2).

TABLEAU I

Poids à 90 jours, des descendants d'une série de onze taureaux testés au C. I. A. de Marmillat

Taureaux	<i>a</i>	<i>b</i>	<i>c</i>	<i>d</i>	<i>e</i>	<i>f</i>	<i>g</i>	<i>h</i>	<i>j</i>	<i>k</i>	<i>l</i>
Nombre de descendants	20	21	48	20	13	16	19	48	23	9	41
Poids moyens des descendants	172,16	162,50	161,54	161,52	161,08	159,74	157,82	154,69	151,37	149,14	149,08

Nous calculons ensuite pour chaque couple de moyennes et en commençant par les moyennes les plus extrêmes, la valeur

$$(\bar{x}_i - \bar{x}_k) \sqrt{\frac{2n_i \cdot n_k}{n_i + n_k}}$$

que nous comparons à la valeur de R_p correspondante.

TABLEAU 2

Valeurs de R_p pour différents intervalles ρ et seuils de probabilité

Intervalle →	2	3	4	5	6	7	8	9	10	11
Seuil de probabilité ↓										
0,10.....	47,61	50,40	52,24	53,61	54,55	55,51	56,21	56,82	57,34	57,81
0,05.....	56,74	59,73	61,76	63,23	64,40	65,36	66,16	66,83	67,43	67,96
0,01.....	74,57	77,70	79,83	81,83	82,70	83,74	84,64	85,40	86,08	86,69

Lorsque deux moyennes ne sont pas significativement différentes, on peut considérer que toutes les moyennes qu'elles encadrent forment un groupe statistiquement homogène au seuil considéré. Ceci n'est réellement vérifié que si les effectifs des groupes sont égaux. L'application du test de DUNCAN dans le cas où le nombre

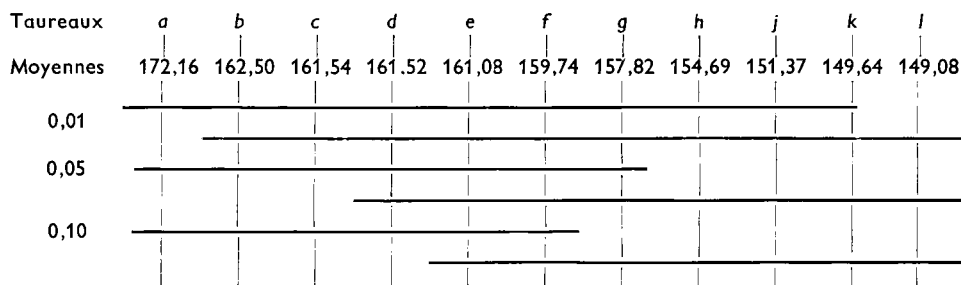
de répétitions est très différent d'une moyenne à l'autre peut nous amener à considérer comme non significatives des différences qui le sont en réalité (tabl. 3).

TABLEAU 3

Calcul des différences significatives entre taureaux

Travaux comparés	p	$(\bar{x}_i - \bar{x}_k) \sqrt{\frac{2n_i n_k}{n_i + n_k}}$	Différences significatives		
			0,10	0,05	0,01
<i>a l</i>	11	119,69	×	×	×
<i>a k</i>	10	79,50	×	×	
<i>a j</i>	9	96,04	×	×	
<i>a h</i>	8	92,77	×	×	
<i>a g</i>	7	63,23	×		
<i>a f</i>	6	52,41			
<i>b l</i>	10	70,72	×	×	
<i>b k</i>	9	45,65			
<i>c l</i>	9	82,85	×	×	
<i>c k</i>	8	46,29			
<i>d l</i>	8	64,56	×		
<i>d k</i>	7	41,82			
<i>e l</i>	7	53,28			

Nous pouvons représenter graphiquement (graph. 1) ces résultats en joignant par un trait continu, les moyennes extrêmes, qui ne sont pas significativement différentes au seuil considéré.



GRAPH. 1

Ainsi dans les résultats de cet exemple qui figurent sur le tableau 3, nous constatons que la différence $\bar{x}_a - \bar{x}_h$ n'est pas significative au seuil de 0,01 à cause du faible nombre d'observations disponibles pour le taureau *k*.

En poursuivant l'analyse, nous trouvons par contre que les différences $\bar{x}_a - \bar{x}_j$ et $\bar{x}_a - \bar{x}_h$ sont significatives au seuil de 0,01.

IV. — NOMBRE DE DESCENDANTS NÉCESSAIRE
POUR FAIRE UN CLASSEMENT PRÉCIS DES TAUREAUX

I. — *Cas général de 2 moyennes*

De façon générale, nous constatons dans l'exemple ci-dessus une différence significative entre les meilleurs taureaux et les plus mauvais ; les autres formant un groupe statistiquement homogène.

Dans l'optique d'un contrôle de descendance et d'une sélection effectués progressivement par le Centre au fur et à mesure que sont obtenues des données sur les descendants de chaque taureau, il serait intéressant de connaître à partir de quel nombre de descendants par taureau, on peut faire un classement précis.

Nous considérerons que nous obtenons un classement précis, lorsque la différence entre 2 moyennes consécutives \bar{x}_i et \bar{x}_{i+1} est significative à un seuil de probabilité fixé à l'avance.

Méthode de calcul.

Dans cette optique, nous avons déterminé le nombre minimum de descendants à contrôler par taureau pour que 2 moyennes consécutives \bar{x}_i \bar{x}_{i+1} soient significativement différentes à un seuil déterminé.

Le nombre de descendants à contrôler par géniteur, peut être estimé en fonction des paramètres caractéristiques de la population étudiée.

σ_p^2 = variance phénotypique du caractère.

h^2 = hérédité du caractère.

$d = \bar{x}_i - \bar{x}_{i+1}$ = différence entre 2 moyennes consécutives,

et de la probabilité P attachée à la connaissance de la différence d ,

Nous poserons en outre :

σ_T^2 = variance intra-taureau = variance due à l'erreur.

σ_G^2 = variance génétique additive.

σ_S^2 = part de la variance génétique additive due au taureau.

Estimation de n.

Dans le cas de 2 moyennes consécutives, le test de DUNCAN donne :

$$(\bar{x}_i - \bar{x}_{i+1}) \sqrt{\frac{2n_i \cdot n_{i+1}}{n_i + n_{i+1}}} > \sigma_T \cdot Z_{2,n}.$$

Si les 2 géniteurs ont le même nombre de descendants, nous aurons :

$$n_i = n_{i+1} = n = \text{nombre de descendants par taureau.}$$

D'autre part, les lots de testage portant toujours sur un grand nombre d'animaux, nous prendrons les valeurs de $Z_{2,n}$ valables pour les grands échantillons.

Nous aurons de ce fait :

$$\begin{aligned} Z_{2,n} &= C^{\text{te}} \text{ pour une probabilité déterminée.} \\ P = 0,01 &\rightarrow Z_{2,n} \longrightarrow = 3,643 \\ P = 0,05 &\rightarrow Z_{2,n} \longrightarrow = 2,772 \\ P = 0,10 &\rightarrow Z_{2,n} \longrightarrow = 2,363. \end{aligned}$$

Dans ce cas, la formule de DUNCAN devient :

$$\begin{aligned} d \sqrt{n} &> \sigma_T \cdot Z_{2,n} \\ \text{d'où} \quad n &> \frac{\sigma_T^2 \cdot Z_{2,n}^2}{d^2}. \end{aligned} \quad (I)$$

Nous devons maintenant exprimer σ_T^2 en fonction des paramètres caractéristiques de la population. En effet, nous savons que :

$$\begin{aligned} \sigma_p^2 &= \sigma_s^2 + \sigma_T^2 \\ \sigma_s^2 &= \frac{1}{4} \cdot \sigma_G^2 \\ \text{d'où} \quad \sigma_T^2 &= \sigma^2 - \frac{1}{4} \sigma_G^2. \end{aligned}$$

D'après la définition de h^2 nous avons :

$$\begin{aligned} \sigma_G^2 &= h^2 \cdot \sigma_p^2 \\ \text{d'où} \quad \sigma_T^2 &= \sigma_p^2 (1 - 0,25 h^2) \end{aligned}$$

En reportant dans l'équation I, nous avons :

$$n > \frac{\sigma_p^2 (1 - 0,25 h^2) Z_{2,n}^2}{d^2} \quad (2)$$

Représentation graphique.

Pour faciliter la représentation graphique des variations de n , nous exprimerons les différences moyennes entre lots consécutifs en unités d'écart type du caractère $\frac{d}{\sigma_p}$,

d'où la formule :

$$n > \frac{(1 - 0,25 h^2) Z_{2,n}^2}{\left(\frac{d}{\sigma_p}\right)^2} \quad (3)$$

Nous pouvons ainsi calculer en nous fixant plusieurs valeurs de h^2 et plusieurs seuils de probabilité, qu'elle est pour différentes valeurs de $\frac{d}{\sigma_p}$ la valeur minimum de n pour que cette différence soit significative au seuil considéré.

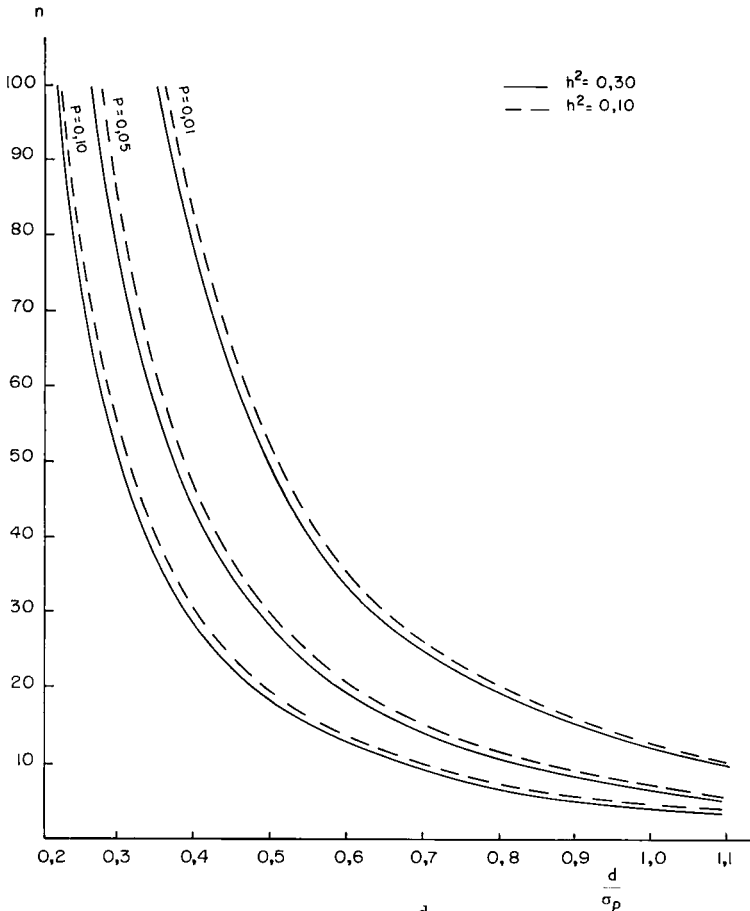
Nous avons effectué ce calcul pour les seuils de probabilités de 0,01 ; 0,05 ; 0,10 et les valeurs d'héritabilité : 0,1 et 0,3 (tabl. 4 et graph. 2).

TABLÉAU 4

Valeurs minimum de n pour que la différence réduite $\frac{d}{\sigma_p}$ soit significative au seuil considéré

Nota : Les chiffres indiqués dans ce tableau sont toujours arrondis à l'unité supérieure

Seuil de probabilité	$h^2 \downarrow$	$\frac{d}{\sigma_p} \rightarrow$										
		0,1	0,2	0,3	0,4	0,5	0,6	0,7	0,8	0,9	1,0	1,2
0,10	$h^2 = 0,3$	463	116	52	29	19	13	10	8	6	5	4
	$h^2 = 0,1$	488	122	55	31	20	14	10	8	7	5	4
0,05	$h^2 = 0,3$	711	178	79	45	29	20	15	12	9	8	5
	$h^2 = 0,1$	746	187	83	47	30	21	16	12	10	8	6
0,01	$h^2 = 0,3$	1 228	307	137	77	50	35	26	20	16	13	9
	$h^2 = 0,1$	1 240	324	144	81	52	36	27	21	16	13	9



GRAPHIQUE 2. — Variation de n en fonction de $\frac{d}{\sigma_p}$ pour différentes valeurs de P et de h^2

Nous constatons que :

— Le nombre d'animaux à contrôler par taureau, varie dans de larges mesures en fonction du seuil de probabilité admis.

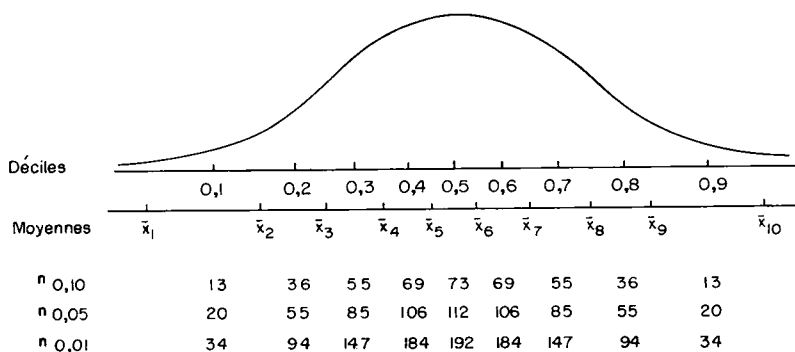
— Pour un seuil de probabilité déterminé, les variations de n entre les 2 valeurs de h^2 considérées sont faibles.

Remarque.

Dans le cas où nous comparons 2 moyennes successives dans une série de moyennes rangées, le test t et le test de DUNCAN, donnent des résultats identiques si le nombre de degrés de liberté correspondant à l'erreur est grand. De ce fait, pour l'étude que nous venons de faire, nous aurions obtenu les mêmes résultats pour $n > 20$ en utilisant le test t ou le test de DUNCAN.

2) Cas de moyennes distribuées normalement

Lorsque nous étudions les résultats se rapportant à un grand nombre de taureaux, nous pouvons nous attendre à ce que les moyennes soient distribuées selon une loi normale. Si nous décomposons la courbe normale en déciles, la probabilité pour une moyenne prise au hasard de tomber dans l'une quelconque des 10 zones ainsi déterminées est la même.



Nous avons donc étudié dans le cas théorique moyen où nous avons 10 moyennes placées chacune au centre d'un intervalle interdécile quel serait le nombre de descendants à contrôler par taureau pour faire un classement précis des géniteurs. Nous avons effectué le calcul pour $h^2 = 0,3$ et pour les probabilités 0,10 ; 0,05 ; 0,01. Les valeurs de n obtenues figurent sur le graphique 3.

Nous constatons que, pour trouver une différence significative au seuil 0,05 entre \bar{x}_1 et \bar{x}_2 (moyennes contiguës les plus extrêmes), il suffit de contrôler 20 descendants par taureau, alors qu'il faut en contrôler 112 pour obtenir une différence significative entre \bar{x}_5 et \bar{x}_6 (moyennes contiguës les plus centrales).

V. — CONCLUSION

La comparaison des performances moyennes relatives aux taureaux appartenant à une série de testage pris 2 à 2 permet d'abord d'observer qu'avec les effectifs de descendants contrôlés pour chaque géniteur, on ne peut différencier statistiquement (aux seuils de 5 p. 100 et 1 p. 100) que les meilleurs taureaux d'un essai d'avec les plus mauvais. Il est possible de vérifier ce résultat sur des séries de taureaux mis en comparaison dans les mêmes conditions de milieu.

Le nombre de descendants à contrôler pour distinguer statistiquement deux taureaux de rangs voisins à un seuil de probabilité donné, est, en moyenne, plus faible pour 2 taureaux dont les moyennes sont situées aux extrémités de la distribution que pour 2 autres dont les moyennes sont centrales. La connaissance de cet « effectif statistique » est intéressante, car les décisions sur le choix des taureaux de testage sont souvent échelonnées dans le temps au fur et à mesure qu'augmente l'information recueillie c'est-à-dire l'effectif de produits contrôlés pour chacun d'eux.

Reçu pour publication en février 1964.

SUMMARY

APPLICATION OF THE TEST OF COMPARISON
OF ORDERED MEANS TO THE PROGENY TEST OF BULLS

DUNCAN's test of comparison of ordered means was applied to the average daughter-performances of artificial insemination bulls.

The method was applied, *a posteriori*, to the results obtained for a number of sires. An *a priori* investigation on the number of offspring needed to show significant differences on the levels of 10 p. 100, 5 p. 100 and 1 p. 100, was also undertaken.

RÉFÉRENCES BIBLIOGRAPHIQUES

- DUNCAN D. B., 1955. Multiple range and multiple F tests. *Biom.*, **11**, 1-42.
HARTER L. H., 1960. Critical values for DUNCAN's new multiple range test. *Biom.*, **16**, 671-685.
KRAMER C. Y., 1956. Extension of multiple range tests to group means with unequal numbers of replications. *Biom.*, **12**, 307-310.
POUTOUS M., VISSAC B., 1962. Recherche théorique des conditions de rentabilité maximum de l'épreuve de descendance des taureaux d'insémination artificielle. *Ann. Zootech.*, **11**, 233-256.
RIVES M., 1959. Sur la comparaison des moyennes dans les essais variétaux. *Ann. Amélior. Plantes*, **3**, 357-376.
ROBERTSON A., 1957. Optimum group size progeny testing and family selection. *Biom.*, **13**, 442-450.