REVUE BELGE DE STATISTIQUE ET DE RECHERCHE OPERATIONNELLE

Vol. 2 - N° 2
JUIN 1961

BELGISCH TIJDSCHRIFT VOOR STATISTIEK EN OPERATIONEEL ONDERZOEK

Vol. 2 - Nr 2
JUNI 1961

La « Revue Belge de Statistique et de Recherche Opérationnelle » est publiée par les Sociétés suivantes :

ABSI (en collaboration avec l'OBAP). — Association Belge pour les Applications Industrielles de la Statistique.
Siège social : 60, rue de la Concorde, Bruxelles 5.
Secrétariat : 10, rue du Tulipier, Bruxelles 19.

AGESCI. — Association Belge pour l'Application des Méthodes scientifiques de Gestion.
Siège social : 4, r. Ravenstein, Bruxelles.
Secrétariat : 2, allée des Platanes, Loverval.

SBS. — Société Belge de Statistique.
Siège social : 44, rue de Louvain, Brux.
Secrétariat : 44, rue de Louvain, Brux.

COMITE DE DIRECTION

S. MORNARD, Ingénieur civil, 51, rue Souveraine, Bruxelles 5.
J. TEGHEM, Professeur à l'U.L.B., 1, av. Reine Marie-Henriette, Bruxelles 19.
J. WANTY, Ingénieur civil, 85, avenue A. Huysmans, Bruxelles 5.

COMITE DE SCREENING

A. HEYVAERT, Licencié en Sciences, 8, av. de la Prospérité, Dilbeek.
Ph. PASSAU, Ingénieur civil, Docteur en Sciences, 2, allée des Platanes, Loverval.
J. TEGHEM, Professeur à l'U.L.B., 1, av. Reine Marie-Henriette, Bruxelles 19.

REDACTION

L. MICHA, Ingénieur civil, 43, rue J.-B. Colijns, Bruxelles 5 - Tél. 44.76.11.

SECRETARIAT

M^me^ J. CROCHELET-GHISLAIN, 10, rue du Tulipier, Bruxelles 19 - Tél. 45.35.29.

Het « Belgisch Tijdschrift voor Statistiek en Operationeel Onderzoek » wordt uitgegeven door de volgende Verenigingen :

ABSI (in samenwerking met B.D.O.P.). — Belgische Vereniging voor Industriële Toepassingen van de Statistiek.
Maatschappelijke zetel : 60, Eendrachtstraat, Brussel 5.
Secretariaat : 10, Tulpenboomstraat, Brussel 19.

AGESCI. — Belgische Vereniging voor de Toepassingen van de Wetenschappelijke Methoden van Bedrijfsbeheer.
Maatschappelijke zetel : 4, Ravensteinstraat, Brussel.
Secretariaat : 2, allée des Platanes, Loverval.

SBS. — Belgische Vereniging voor Statistiek.
Maatschappelijke zetel : 44, Leuvensestraat, Brussel.
Secretariaat : 44, Leuvensestraat, Brussel

DIRECTIE COMITE

S. MORNARD, Burgerlijk Ingenieur, 51, rue Souveraine, Brussel 5.
J. TEGHEM, Professor aan de V.U.B., 1, Koningin Marie-Henriettelaan, Brussel 19.
J. WANTY, Burgerlijk Ingenieur, 85, A. Huysmanslaan, Brussel. 5.

SCREENING COMITE

A. HEYVAERT, Licenciaat in de Wetenschappen, 8, av. de la Prospérité, Dilbeek
Ph. PASSAU, Burgerlijk Ingenieur - Dr. in de Wetenschappen, 2, allée des Platanes, Loverval.
J. TEGHEM, Professor aan de V.U.B., 1, Koningin Marie-Henriettelaan, Brussel 19.

REDACTIE

L. MICHA, Burgerlijk Ingenieur, 43, J.-B. Colijnsstraat, Brussel 5 - Tel. 44.76.11.

SECRETARIAAT

Mevr. J. CROCHELET-GHISLAIN, 10, Tulpenboomstr. Brussel 19. Tel. 45.35.29.

Revue belge de Statistique et de Recherche opérationnelle

VOL. 2 — N° 2 — JUIN 1961 VOL. 2 — Nr 2 — JUNI 1961

SOMMAIRE ——————— INHOUD

A. HEYVAERT Editorial.
Van de Redactie.

J.M. DOPCHIE Les groupes de travail de l'AGESCI.

P. DE MUNTER Méthodes statistiques pour la comparaison de plus de deux traitements. I.

Cycle de cours sur le contrôle de réception par échantillonnage.

Lessencyclus over afnamecontrole.

Informations.

Mededelingen.

Belgisch Tijdschrift voor Statistiek en operationeel Onderzoek

EDITORIAL

Possibilites et limitations des techniques statistiques

1. — Il est devenu une habitude dans le monde moderne de vouloir tout définir. Quoique ce désir de précision soit très louable en soi, il est cause de nombreuses difficultés dès que nous essayons de « définir » une groupe de techniques encore en pleine évolution comme c'est le cas dans le domaine « statistiques ». Comme il n'entre pas dans notre intention d'ouvrir ici de nouvelles polémiques à ce sujet, nous nous contentons de constater que le terme

« techniques statistiques »

couvre généralement un ensemble d'outils, apparemment très hétérogène (voyez par exemple la théorie des moindres carrés et l'analyse séquentielle), mais tous gérés par les mêmes principes de base; les lois du hasard ou des probabilités.

2. — Et déjà nous rencontrons, dans cette constatation même une première limitation : en effet, le hasard (encore lui !) a lancé l'étude de ces lois du hasard entre les mains de mathématiciens : quoi d'étonnant, dans ces conditions, que les outils sont devenus mathématiques et que, par conséquent, l'utilisation des techniques statistiques est restée longtemps limitée aux seules possibilités mathématiques. Il n'y a que depuis quelques années que les techniques de simulation — genre Monte Carlo et autres — ont acquis le droit de cité dans notre arsenal d'outils !

3. — Il n'y a pas tellement longtemps les techniques statistiques étaient encore peu connues et peu utilisées. Le succès foudroyant des premières applications — c'est souvent le cas pour une nouveauté — a eu comme conséquence de faire naître le mythe de l'outil magique qui va tout résoudre, la fée aux possibilités illimitées...

Insuffisamment ou mal informé, l'utilisateur en puissance applique cette merveille à un problème dont la solution dépasse les possibilités réelles de l'outil... l'échec suit et presque toujours la conclusion très humaine attaque le pauvre outil qui ne saurait se défendre :

« J'ai échoué, donc l'outil ne vaut rien ».

Nous pensons vraiment que pour éliminer la possibilité d'une telle conclusion, la description d'un quelconque outil devrait toujours *être précédée par celle de ses limitation !*

4. — Une nouvelle limitation est inhérente à la nature même des techniques statistiques qui se servent, comme matières premières, des données d'observations :

Il est absolument impossible de tirer des données, des informations qui ne s'y trouvent pas.

En d'autres termes :

— *l'absence de données,*

— *des données inexactes,*

— *des données arbitrairement sélectionnées,*

— *des données enregistrées à la légère,*

ne peuvent pas permettre d'arriver à des conclusions valables.

Il en résulte qu'un statisticien digne de ce nom attachera toujours *une grande importance à la collecte et à la vérification des données.*

5. — L'application des techniques statistiques, apparemment facile, par des personnes non suffisamment initiées, de même que l'interprétation tendancieuse des données comme des résultats peut permettre d'arriver à des conclusions fausses.

Ce qui nous permet de répéter que les statistiques constituent un moyen scientifique pour mentir.

6. — Enfin, comme les statistiques ne constituent pas un but en soi, leur application ne saurait être efficace qu'à condition de connaître parfaitement le problème à résoudre. Nous pensons qu'il est prudent de ne pas sousestimer cette limitation, source de nombreux déboires :

Il ne suffit pas d'être statisticien, il faut encore parfaitement connaître le problème à étudier !

7. — Quelles que sévères que soient les limitations, elles ne diminuent en rien les magnifiques possibilités des techniques statistiques, sainement appliquées.

L'examen que nous venons de faire de ces mêmes limitations nous permet d'ailleurs de situer exactement le domaine de ces immenses possibilités :

Les techniques statistiques constituent un puissant outil d'investigation partout où existent — ou peuvent être obtenues — des données numériques.

Il n'entre pas dans le cadre de cet éditorial de passer en revue toutes les techniques ni toutes leurs possibilités. Nous nous limitons volontairement à quelques-unes des plus grandes.

8. — La connaissance d'un événement n'étant que très exceptionnellement totale *(et le plus souvent même très fractionnaire) les techniques statistiques nous donnent l'inestimable possibilité :*

d'apprendre à connaître — avec un degré de précision dépendant naturellement du problème et des données, mais souvent étonnement élevé — les caractéristiques d'un ensemble à partir des caractéristiques connues d'une de ses fractions.

Les techniques d'échantillonage illustrent bien cette possibilité;

l'analyse de variance, les calculs de régression et de corrélation entrent dans ce domaine.

9. — Un événement évoluant souvent dans le temps, *son évolution future entre automatiquement dans la fraction ignorée de la connaissance totale.*

L'importance qu'a eu de tous temps — et qu'a de plus en plus — pour la conduite des entreprises, une certaine connaissance a priori de cet avenir (ce ne sera d'ailleurs jamais qu'une estimation plus ou moins probable) permet d'estimer à sa juste valeur l'apport des techniques statistiques qui permettent d'augmenter — souvent d'une façon étonnante — la précision de ces estimations (statistiquement nous dirons : réduire les limites de l'incertitude).

Les cartes de contrôle et les séries chronologiques entrent dans ce domaine.

10. — Enfin, refusant de se soumettre au seul hasard pour lui obtenir les données numériques nécessaires à son application, la statistique a créé une nouvelle possibilité :

prédéterminer les données numériques nécessaires pour obtenir le meilleur rendement de l'application des autres techniques statistiques.

Ainsi la connaissance du hasard permet d'éliminer en partie ce même hasard dans le domaine de la recherche par l'étude des plans d'expérimentation.

11. — Nous ne pensons pas avoir passé en revue ni l'ensemble des limitations, ni surtout l'ensemble des possibilités offertes par les techniques statistiques, mais nous espérons par l'examen de ces quelques points majeurs avoir mis en lumière les véritables possibilités — immenses croyons-nous — qu'offrent les techniques statistiques.

A. HEYVAERT,

Directeur à l'I.O.I.C.

VAN DE REDACTIE

Toepassingsmogelijkheden en beperkingen bij het gebruiken van statistische technieken

1. — Het is een gewoonte geworden in onze moderne wereld alles te willen bepalen. Alhoewel deze neiging op zichzelf genomen goed is, veroorzaakt ze nochtans vele moeilijkheden zodra wij beproeven een bepaling te vinden voor een groep technieken welk nog in volle ontwikkeling is zoals de statistieken. Daar het nu niet in onze bedoeling ligt een nieuwe polemiek te openen, stellen we ons tevreden met de vaststelling dat de term

« statistische technieken »

over het algemeen een stel werktuigen dekt, ogenschijnlijk zeer uiteenlopend, (zie b.v.b. de theorie der minimale kwadratische afwijkingen en de sekwentiële analyse), maar allen door éénzelfde basisprincipe beheerd :

de wetten van het toeval of de kansrekening

2. — Deze vaststelling geeft reeds aanleiding tot het ontdekken van een eerste beperking :

inderdaad, het toeval (weeral !) heeft het zo gewild dat haar wetten door wiskundigen werden bestudeerd.

Het is dan ook niet te verwonderen dat de gebouwde werktuigen wiskundig zijn en dat de statistische mogelijkheden dan ook lang beperkt bleven tot die der wiskunde. Het is inderdaad nog niet zolang dat simulatietechnieken (zoals Monte Carlo b.v.b.) burgerrecht verkregen in ons werktuigarsenaal.

3. — In een nabij verleden waren statistische technieken slechts weinig gekend en weinig gebruikt. Het schitterend welslagen der eerste toepassingen (dit is dikwijls het geval voor een nieuwigheid) gaf ongelukkig aanleiding tot de geboorte van een mythe... een magisch werktuig dat alles oplossen kan, een fee met onbeperkte mogelijkheden...

Onvoldoende of slecht ingelichte personen pasten dan dit wonder-

middel toe op een probleem dat ver buiten de werkelijke toepassingsmogelijkheden valt.

En het onvermijdelijke falen wordt natuurlijk — een echt menselijk besluit — ten laste gelegd van het weerloze werktuig :

« *Ik faalde,* dus deugt het werktuig niet ! »

Wil men nu werkelijk deze fout uitschakelen dan is het noodzakelijk de beschrijving der mogelijkheden te doen voorafgaan door de opsomming der beperkingen !

4. — Een andere beperking is innig verwand aan de intieme aard zelf der statistieken welke als grondstof waarnemingsgegevens *gebruiken :*

het is absoluut onmogelijk meer uit deze gegevens te trekken dan er werkelijk inzit.

Anders gezegd :

— *het ontbreken van gegevens,*
— *onnauwkeurige gegevens,*
— *arbitrair uitgekozen gegevens,*
— *lichtzinnig opgenomen gegevens,*

kunnen niet tot geldige besluiten leiden.

Hieruit volgt dan ook dat iedere waardige statistieker altijd *een groot belang hecht aan het verkrijgen en het nazien der gegevens.*

5. — Het aanwenden van statistische technieken, ogenschijnlijk gemakkelijk, door onvoldoende ingeweide personen, evenals het tendensieuze uitleggen van gegevens of berekeningsresultaten kan natuurlijk aanleiding geven tot valse besluiten.

Dit laat ons toe hier te herhalen dat statistieken een wetenschappelijke wijze van liegen kunnen worden.

6. — En daar de statistieken nu eenmaal niet een doel, maar wel een middel zijn, kan hun toepassing alleen doelmatig zijn wanneer het bestudeerde vraagstuk grondig gekend is. Wij zijn de mening toegedaan dat deze beperking welke niet te onderschatten is, de bron is van talrijke ontgoochelingen :

het is niet voldoende statistieker te zijn, maar men moet tevens grondig het te bestuderen probleem kennen.

7. — Hoe streng de beperkingen nu ook wezen, ze verminderen geenszins de buitengewone mogelijkheden der gezond toegepaste statistische technieken.

Wat we tot nu toe zegden laat ons trouwens toe nauwkeurig het domein dezer mogelijkheden te beschrijven.

De statistieken vormen een machtig onderzoekstuig overal waar numerieke gegevens bestaan (af kunnen verkregen worden).

Het valt niet in onze bedoeling hier nu alle technieken of alle mogelijkheden te bespreken. Wij beperken ons enkel tot de belangrijkste.

8. — Een gebeuren is ons slechts heel zelden volledig bekend *(gewoonlijk zelfs slechts heel gedeeltelijk !). Het is dan ook een waardevolle mogelijkheid welke de statistieken ons brengen :*

de karakteristieken van een gebeuren — met een benadering

welke natuurlijk van het probleem en van de gegevens afhangt — afleiden van de gekende karakteristieken een zijner frakties.

De steekproeftheoriën belichten duidelijk deze mogelijkheid; de variantie analyse, de regressie en korellatieberekening vallen ook binnen dit domein.

9. — Een gebeuren evolueert dikwijls met de tijd, zijn toekomstig gedragen valt automatisch binnen het onbekende deel van de « totale » kennis. Het steeds groter wordende belang voor het zakenleven, van een zekere « a priori » kennis dezer toekomst laat toe de door de statistieken gebrachte mogelijkheid naar haar juiste waarde te schatten :

deze technieken laten toe een nauwkeuriger beeld te krijgen over deze toekomst (statistisch zouden we zeggen : de grenzen van het onzekere meer beperken).

Kontrolekaarten en Tijdreeksen vallen binnen dit domein.

10. — Uiteindelijk weigerden de statistieken het bekomen der nodige gegevens alleen aan het toeval over te laten en bouwden een nieuwe mogelijkheid :

vooraf bepalen welke gegevens nodig zijn teneinde het beste rendement te verkrijgen bij het toepassen der statistieken.

Zodoende laat de kennis van het toeval toe ditzelfde toeval gedeeltelijk uit te schakelen door het uitwerken van de

Proevenschema's.

11. — Wij beweren natuurlijk niet in deze bondige nota noch de beperkingen, noch de mogelijkheden der statistische technieken volledig te hebben overlopen, maar we hopen wel dat het korte overzicht van sommige hoofdpunten ervan ons toeliet de buitengewone mogelijkheden dezer technieken te belichten.

A. HEYVAERT,
Direkteur I.O.I.C.

Les groupes de travail de l'Agesci

par J. M. DOPCHIE.

En vue de diffuser parmi les membres de l'AGESCI les connaissances acquises en matière de méthodes scientifiques de gestion, et de les faire participer à l'activité des groupes de travail, le Comité de Rédaction de la Revue Belge de Statistique et de Recherche Opérationnelle a décidé de publier dans sa Revue des compte rendus assez larges de leurs travaux.

Lorsque les groupes ont été formés, il ne s'agissait pas d'y donner un enseignement postuniversitaire sur l'une ou l'autre technique mais bien d'une prise de conscience des problèmes de gestion en favorisant les échanges d'expériences et en faisant ressortir les difficultés rencontrées dans la résolution de ces problèmes. Plutôt que de faire appel à la méthode des cas, on a préféré demander à chaque participant de faire part de ses préoccupations pour élaborer ensuite sous la direction d'un conseiller qualifié un modèle synthétique abordant les divers aspects sous lesquels les problèmes se posent à l'entreprise.

Afin d'homogénéiser le groupe il a fallu faire quelques exposés de théorie et d'autre part des cas pratiques ont illustré l'intérêt de l'approche scientifique de certains problèmes.

Dans ce numéro, nous commençons par le groupe « Gestion du Matériel » qui a été constitué le premier par l'AGESCI, au mois de mai 1959, et qui a travaillé suivant cette méthode sous la direction de Monsieur A. KAUFMANN.

De nederlandse tekst van het artikel hieronder zal in ons tijdschrift nr 3 verschijnen.

GESTION DU MATERIEL

Introduction : Exposé des problèmes.

Cette rubrique comporte une très grande variété de problèmes : l'usure - la maintenance - le remplacement - les investissements. De plus, l'entreprise, constituant un tout, il est difficile de les traiter sans aborder ceux des stocks, de la programmation et de la prévision. Cependant ceux-ci faisant l'objet de l'activité d'autres groupes, on s'efforça de les négliger pour concentrer les discussions sur la maintenance qui constitue le point primordial d'une gestion de matériel.

Dans une première entreprise il s'agit principalement de la détermination d'une politique de remplacement. L'atelier de production considéré comporte un nombre considérable de machines mises en parallèle ayant un taux de production élevé. Ce parc ayant un âge assez avancé, la direction de cette entreprise désirait mettre en évidence un ordre de priorité de remplacement et la série optimum de mise en fabrication des nouvelles machines. Dès le premier exposé on s'est aperçu qu'il était nécessaire d'avoir un langage commun et qu'il fallait établir quelques définitions.

Par une politique on entend le choix effectué sur un paramètre d'une structure dont on a le contrôle. Lorsque dans un problème de gestion plusieurs politiques sont possibles, le but de la recherche est la détermination de la politique qui optimisera le résultat final de l'opération. Ce résultat se décrit sous la forme d'une fonction dite économique qui est unique. Cette fonction économique comporte dans le cas traité des coûts d'investissements et des coûts d'avarie ; ces derniers comprennent les frais de stockage des pièces de rechange, les pertes de production, les frais de réparation.

Dans une seconde entreprise, on considère des machines mises en série, où tout arrêt de l'une d'elle entraîne l'arrêt de toute la chaîne. Le problème posé est celui d'une politique d'entretien. L'entretien est l'ensemble des opérations que l'on doit réaliser pour maintenir un matériel en état de produire une certaine quantité dans une certaine qualité.

Faire les dépenses qui consistent à réaliser cette performance exactement, c'est de l'entretien juste. Les dépenses que l'on fait au delà, c'est du sur-entretien ; en deçà, c'est

du sous-entretien. Il faut rechercher le moyen de descendre d'un sur-entretien pour se rapprocher autant que possible de l'entretien juste.

La troisième entreprise dispose d'un parc très étendu de véhicules automobiles. Plusieurs problèmes ont été soulevés : l'entretien accidentel - l'entretien systématique - la refonte du véhicule - son remplacement - la localisation et l'importance des ateliers de réparation - le nombre de véhicules mis en service.

Il n'était pas possible d'en examiner tous les aspects soulevés. Ce seront principalement l'entretien et le remplacement qui seront considérés au cours des séances suivantes.

Dans la quatrième entreprise, il existe d'abord un problème de stocks. Le matériel considéré comporte un nombre d'articles de l'ordre de 150.000. Cependant on y trouve également un problème d'investissements pour un matériel dont l'obsolescence est très rapide.

L'exposé de ce cas a fait ressortir les difficultés de mettre au point des critères de sélection auxquelles s'ajoute encore des difficultés psychologiques qui souvent éloignent la décision de l'optimum économique.

La cinquième entreprise dispose d'un nombre étendu de données. Son objectif est de réduire le prix de revient sans augmenter le risque d'avaries. La révision des équipements portant sur plus de 50 % des dépenses relatives au matériel, c'est surtout sur l'entretien préventif et le reconditionnement de certains appareils que s'est penchée cette entreprise.

Dans la sixième entreprise, c'est encore le reconditionnement qui constitue le problème étudié. Contrairement au cas précédent, le matériel est de peu de valeur ; le coût du reconditionnement n'atteint que 12 % de la valeur unitaire mais l'opération s'avère rentable pour un équipement comportant près de 625.000 appareils. Il s'agit ici de déterminer la politique de reconditionnement à suivre.

Au contraire dans la dernière entreprise, il s'agit d'un problème d'investissement : qui revient à un choix entre une modification portée à du matériel existant et l'acquisition d'un équipement neuf. Ceci se rapportait à un parc de 120 machines groupées par séries de 3 à 5 unités.

§ 1er. — *Aspects théoriques.*

Tous ces problèmes de gestion du matériel sont liés à la maintenance. Un matériel de production — (dans une compagnie de transport le matériel produit des tonnes/km ou des voyageurs/km) est sujet à des avaries et à l'obsolescence.

Considérant d'abord les avaries, il est intéressant d'établir ce qu'on appelle « la courbe de survie » des pièces, ou éventuellement des machines, sujettes aux avaries. La courbe de survie représente la variation du nombre d'équipements en service en fonction du temps. Si au temps O on part avec un matériel composé de n_o unités, au temps t il n'y aura plus que n unités en opération.

La probabilité pour qu'un matériel atteigne un âge donné t est donné par le rapport n/n_o.

L'allure de la courbe de survie décèle immédiatement le degré d'homogénéité du matériel.

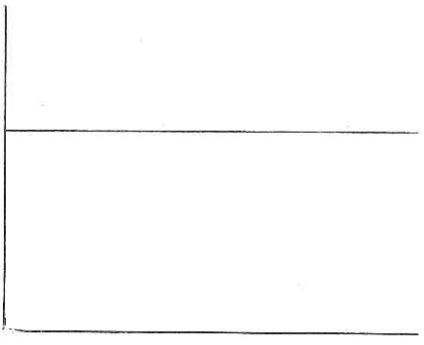

FIG. 1 a.

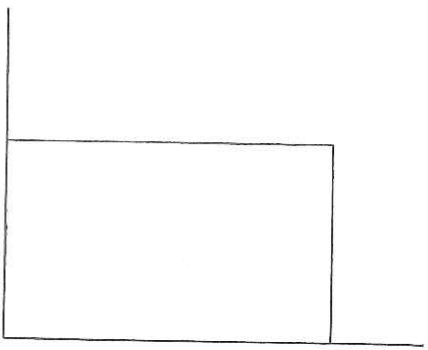

FIG. 1 b.

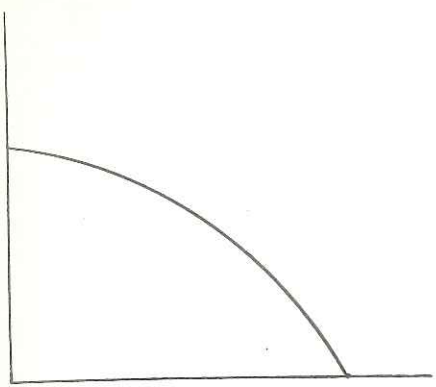

FIG. 1 c.

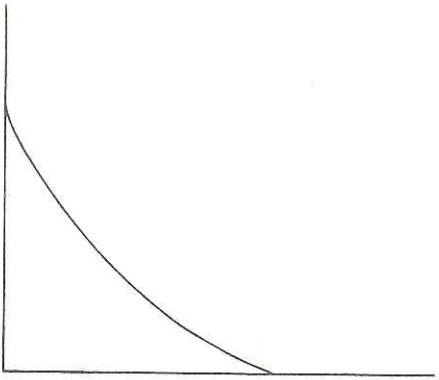

FIG. 1 d.

La courbe 1*a* représente un ensemble d'équipements homogène et inusable.

La courbe 1*b* représente un ensemble d'équipements homogène qui périt en une fois.

La courbe 1*c* appartient à un (matériel) ensemble d'équipements homogène, tandis que la courbe 1*d* est celle d'un matériel peu homogène.

Pour établir la courbe de survie par des moyens statistiques, on doit nécessairement connaître l'histoire du matériel pendant une période de temps, la plus longue possible. On ne dispose pas toujours des données nécessaires bien qu'il apparaît de plus en plus qu'une collecte justicieuse de données valables soit hautement rentable. Si les données manquent on fera une hypothèse basée sur l'examen de quelques cas que l'on a pu recueillir, quitte à traiter le problème par simulation en ajustant la courbe de survie au fur et à mesure.

Dans une étude publiée dans la Revue Française de Recherche Opérationnelle (1959 - premier trimestre), Jacques Kelly examine la justification de l'entretien préventif. Des essais effectués, il semble qu'avec des paramètres différents une répartition lognormale représente assez bien la durée de vie des pièces, telles que lampes, courroies, roulements à bille.

La probabilité $p(t)$ pour qu'une pièce de durée médiane V (l'âge atteint par la moitié des pièces) casse avant l'âge t est donnée par

$$\frac{1}{\sqrt{2\,\pi}} \int_{-\infty}^{x} e^{-\frac{x^2}{2}}\, dx$$

où x est une variable auxiliaire $= \dfrac{\mathrm{Ln}\,(t/\mathrm{V})}{\lambda}$ (Ln = logarithme népérien)

c.à.d. la courbe de survie est définie par deux paramètres V, la durée médiane de la pièce, et λ, un coefficient de dispersion.

La durée moyenne de vie $\bar{t}$, ou l'abscisse partageant en deux l'aire comprise entre la courbe de survie et l'axe des t est donnée par :

$$\bar{t} = \mathrm{V}\, e^{\frac{\lambda^2}{2}}$$

Lorsque le coefficient de dispersion λ est petit, la distribution lognormale se rapproche de la distribution normale et λ tend vers σ/V, écart type relatif.

Suivant Kelly $\lambda = 0{,}3$ s'applique parfaitement à des pièces mises hors service par usure tandis que $\lambda = 0{,}9$ paraît convenir pour des pièces qui périssent par fatigue, bien que parfois on utilise pour ces pièces, la distribution de Poisson assez voisine.

La rupture d'une pièce, de quelque type qu'elle soit, entraîne deux types de dépenses : le coût de la réparation d et le coût des conséquences D (perte de production, etc.).

Dans une première politique, où l'on ne fait pas d'entretien préventif, le coût horaire dû aux ruptures des pièces considérées s'établit à

$$C_1 = \frac{d_1 + D_1}{\overline{t}}.$$

Dans une deuxième politique, où l'on remplace la pièce à l'âge θ, le coût horaire s'établit à

$$C_2 = \frac{d_2 + [1 - p(\theta)]\, D_2}{\overline{t\,\theta}}$$

où $p(\theta)$ est la probabilité pour que la pièce atteigne le temps θ ou encore ne casse pas avant θ et $[1 - p(\theta)]$, son complément, la probabilité pour que la pièce casse avant θ.

L'entretien préventif sera justifié si $C_2 < C_1$ ou $C_2/C_1 < 1$. Dans ce cas on cherchera l'âge θ qui correspond au minimum de C_2/C_1.

Prenant les dépenses de réparation d et le coût D des conséquences, égaux dans les deux politiques, Kelly détermine une limite pratique d'entretien préventif justifié dans un diagramme $(d/D, \lambda)$. (Fig. 1.5).

On fait de l'entretien préventif pour des pièces d'usure, quand pour un λ variant de 0,3 à 0,6, le rapport d/D ne dépasse pas 0,8 à 0,2. Pour les pièces de fatigue, qui sont souvent très coûteuses, il est très rare que l'entretien préventif se justifie. Ce n'est que pour des valeurs de d/D inférieures à 0,05. D'ailleurs, si l'on devait admettre comme courbe de survie la loi de Poisson, le risque de casse dans la période suivant le remplacement est le même pour la pièce neuve que pour la pièce ayant un certain âge.

Dans la construction d'une machine où interviennent plusieurs pièces d'usure, il y a intérêt à les dimensionner de manière à obtenir des caractéristiques de survie analogues afin de pouvoir fixer une périodicité des révisions.

Souvent les dépenses de réparation et les conséquences de l'arrêt ne sont pas les mêmes dans les deux politiques envisagées.

Il suffit de considérer la valeur de la perte de production. La perte de production est généralement plus coûteuse lors d'une avarie que pendant un arrêt pour un entretien préventif. La valeur d'une perte de production est souvent discutée. On peut prendre comme bases la marge bénéficiaire, la valeur de remplacement du produit fini ou la valeur de la matière première, ces valeurs pouvant être multipliées par un certain taux de pénurie, si doivent intervenir des facteurs commerciaux ou de sécurité civile ou militaire.

Il existe donc plusieurs valeurs et c'est pourquoi il y a intérêt de faire fixer le coût de pénurie par la Direction Générale, ou de commun accord par les participants de l'étude. Comme ceux-ci sont souvent de discipline différente on peut utiliser la méthode du vote pondéré. Chaque participant donne un certain poids à chaque paramètre intervenant dans le problème et l'on admet la moyenne des poids comme représentative de l'ensemble des tendances ; pour l'information des participants de l'étude, une étude analytique paramétrée du coût de pénurie doit souvent être entreprise.

Quant à la dépréciation et à l'obsolescence du matériel, ce sont des notions qui tout en étant bien claires à l'esprit ne se valorisent pas aisément. Se référant aux ouvrages de P. Massé : « Le choix des Investissements » et de J. Lesourne : « Technique économique et Gestion industrielle » on peut facilement appliquer ce qui y est écrit sur les investissements en général au matériel en particulier.

On y retrouve la notion d'actualisation qui consiste à poser l'équivalence entre une suite de revenus (ou de dépenses) $R_0, R_1, R_2 \ldots\ldots R_n$ pour les n années futures et la valeur actuelle qu'ils représentent.

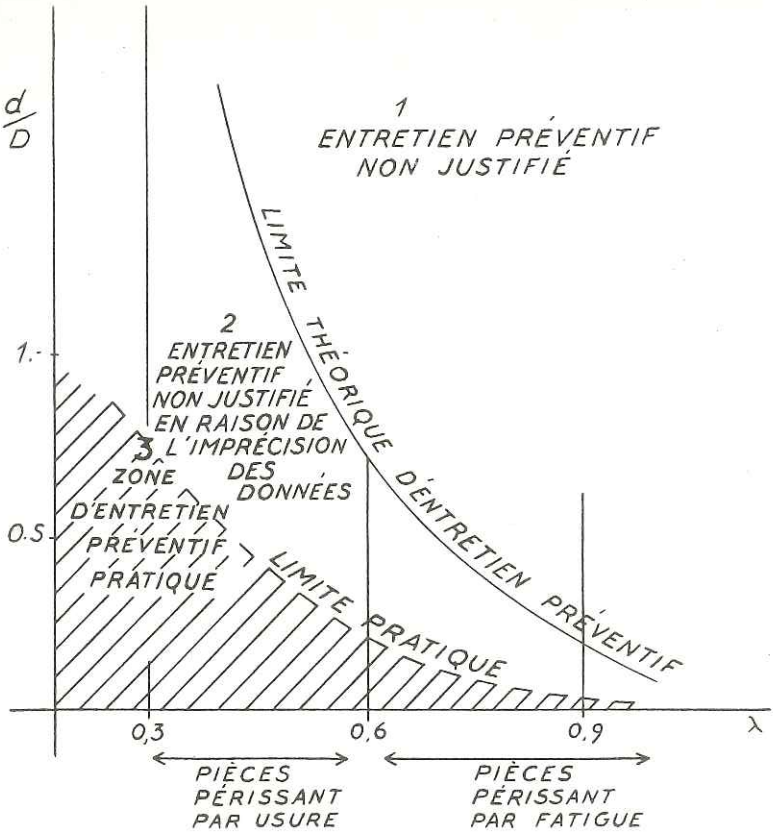

FIG. 1.5.

$$B = R_0 + \frac{R_1}{1+i_1} + \frac{R_2}{(1+i_1)(1+i_2)} + \ldots + \frac{R_n}{(1+i_1)(1+i_2)\ldots(1+i_n)}$$

La valeur d'un matériel est la somme des valeurs actualisées de ses revenus nets futurs. La « juste » dépréciation est ce qu'il faut ajouter à cette valeur du moment pour retrouver la valeur à neuf. Malheureusement cette valeur du moment ou du marché, n'est connue que dans quelques cas particuliers. C'est par exemple celui de l'automobile, où le Moniteur de l'Automobile fournit la valeur désirée. Monsieur Kaufmann a d'ailleurs traité ce problème dans son livre. « Méthodes et Modèles de la Recherche Opérationnelle ». Souvent il faut estimer au lieu de constater. Pour des raisons pratiques on utilise assez souvent comme charge financière d'une année, la somme de la n^{me} partie de la valeur initiale et de l'intérêt de la valeur non encore amortie du matériel.

L'obsolescence peut se décrire sous une forme d'un risque, c'est-à-dire d'une probabilité qu'apparaisse une machine d'un nouveau modèle plus perfectionné. Cette probabilité croît avec le temps, et la courbe a une allure Gaussienne. Dans le modèle synthétique, on admettra que c'est une courbe de Gauss, dont la dispersion constituera un des paramètres du modèle.

Cependant avant d'aborder l'élaboration du modèle, il a été jugé utile de faire un commentaire sur certaines lois des probabilités, lois de distributions et tests de significations.

Il est en effet très important de pouvoir contrôler les hypothèses que souvent on est amené à poser dans l'étude d'un problème.

Ont été traités : Les chaînes de Markov — Les équations de Chapman — Kolmogorov — Les chaînes stochastiques non Markoviennes — Le processus de Poisson — La distribution d'Erlang — La distribution hyperexponentielle ou de Morse — Les tests en X^2 (Pearson) — en t (Student) — en F (Fisher).

§ 2. — ETABLISSEMENT DU MODELE.

Le modèle proposé n'est pas un modèle omnibus permettant l'étude directe des problèmes réels intéressant les membres du groupe, mais il est censé contenir une association des difficultés les plus fréquemment rencontrées et que l'on a cherché à mettre en évidence. Il doit être assez simple pour ne pas exiger des séances de travail trop longues ou trop nombreuses mais il doit éviter d'être caricatural et sa raison d'être est pédagogique ; en fait il prétend constituer un « ensemble de travaux pratiques de gestion scientifique » à l'instar des travaux de physique dans une faculté des Sciences.

2.1. — *Enoncé du modèle.*

Un processus de fabrication comprend 3 chaînes numérotées (1) — (2) — (3) — voir fig. 2.1.

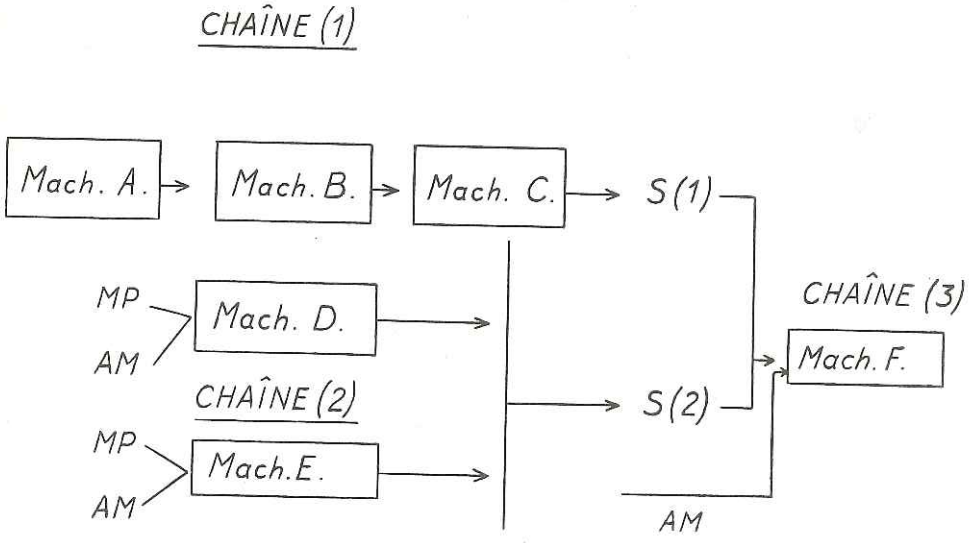

FIG. 2.1.

La chaîne (1) comprend 3 machines A, B. C en cascade, l'arrêt d'une machine entraîne l'arrêt de la chaîne. La chaîne 2 comprend 2 machines D et E en parallèle qui réalisent un même produit ; l'arrêt de D n'entraîne pas l'arrêt de E et réciproquement, mais l'arrêt d'une machine diminue la production de la chaîne (2) de moitié, par unité de temps.

Les produits (1) et (2) sont assemblés dans la chaîne (3) comprenant une machine unique F, on entrevoit des stocks tampons S (1) et S (2) avant F. Les machines A, B et C sont différentes, les machines D et E sont identiques.

2.2. — *Hypothèses générales.*

2.2.1. — Toutes les pannes sont supposées indépendantes les unes des autres. Les machines contiennent différentes pièces sensibles, les unes à la fatigue, les autres à l'usure. On admettra que les unes et les autres de ces pièces ont une courbe de survie selon une loi lognormale de paramètre de dispersion λ ($\lambda = 0{,}3$ pour les pièces d'usure et $\lambda = 0{,}9$ pour les pièces de fatigue).

En plus des pannes principales provoquées par ces pièces, il existe sur chaque machine une source de petites pannes qui se produisent selon une loi de Poisson de taux a.

2.2.2. — On admet que les pannes des machines constituent un phénomène sensiblement non stationnaire en ce qui concerne l'usure et la fatigue. Si $V_m(0)$ est la durée de vie moyenne à $t = 0$, cette durée de vie au temps t n'est plus que $V_m(t) = \varphi(t)\, V_m(0)$.

$\varphi(t)$ représente le vieillissement de la machine et se traduit par une courbe dite de vieillissement fig. 2.2.

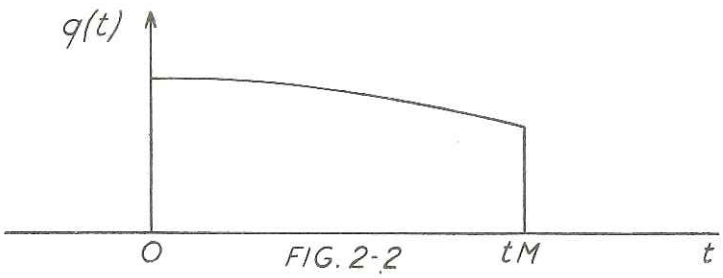

FIG. 2-2

On admet qu'au temps t_M il faut changer impérativement la machine.

2.2.3. — Pour chaque machine on estime une courbe de dépréciation (fig. 2.3).

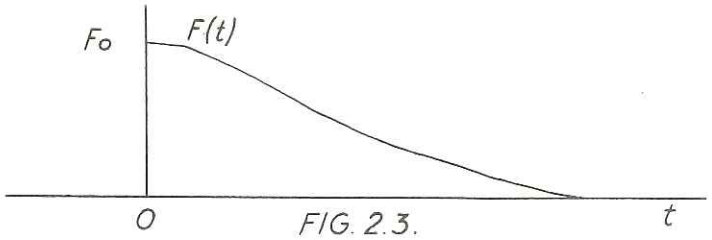

FIG. 2.3.

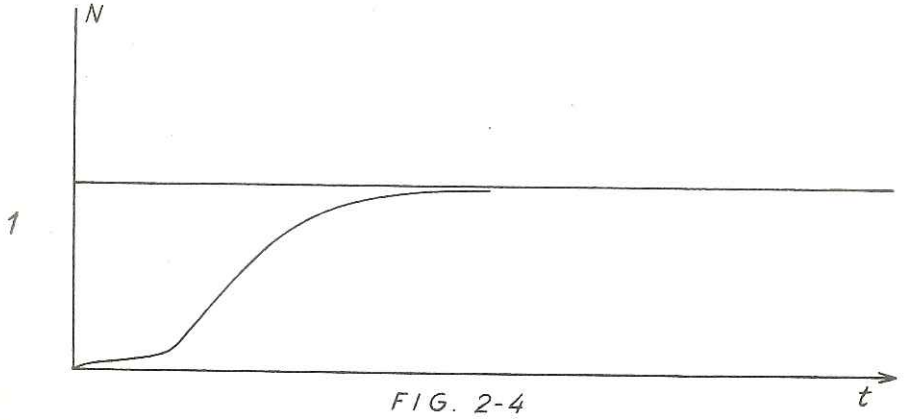

FIG. 2-4

On admet que le prix de revente suit une allure exponentielle ou lognormale. En appelant $F(t)$ la valeur à la période t et F_o la valeur à l'achat on tiendra compte de la valeur instantanée actualisée de chaque machine.

2.2.4. — Cette valeur actuelle est affectée par l'obsolescence provoquée par la création sur le marché d'un type nouveau de machine. L'obsolescence agit sur le prix de revente si une nouvelle machine est découverte, on admet que $F(t)$ doit être multiplié par un coefficient α en %.

La valeur actuelle devient $F'(t) = \alpha F(t)$. La valeur découverte d'une nouvelle machine et sa proposition sur le marché est un phénomène assez rare qui sera obtenu par tirage du sort sur une courbe d'obsolescence à toutes les périodes. Cette courbe d'obsolescence croît avec le temps (fig. 2.4).

Comme on ne connaît pas cette courbe dont l'allure est Gaussienne, on admet que c'est une courbe de Gauss ayant une dispersion qui constitue un des paramètres du modèle.

2.2.5. — Tous les coûts sont actualisés à un taux de l'argent r, supposé constant sur l'ensemble des périodes considérées. L'actualisation est faite en fin d'année. Si l'on se donne un intervalle de description de n périodes et si l'on appelle $f(r)$ les dépenses de fonctionnement et d'entretien pour l'année n, celles-ci interviennent avec leur valeur actualisée

$$\frac{f(r)}{(1+r)^n}.$$

2.3. — *La simulation.*

Le programme des travaux se fait comme devant aboutir à une simulation au moyen d'un calculateur électronique utilisé pour tester les diverses politiques (simulation Machines M). Eventuellement si des décisions intermédiaires devaient intervenir dans les séquences d'une politique, on pourrait concevoir l'étude du problème par une simulation hommes-machines HM destinée à reconstituer avec une certaine fidélité l'atmosphère intime de l'entreprise et à étudier les circuits de décisions ou d'opérations. Il ne s'agirait plus d'un jeu d'entreprises — interentreprise — mais intra-entreprise permettant au groupe d'assister, en temps contracté, à l'évolution, à la mise en service ou à l'obsolescence, aux pannes et autres incidents qui interviennent dans la gestion du matériel d'une entreprise.

2.3.1. — *Le temps.*

On examinera le résultat de chaque politique après 5 ans d'activité ou horizon économique. L'Unité de temps choisie sera l'heure, l'actualisation sera calculée chaque année. L'année comprendra 48 semaines, une interruption de 4 semaines étant admise, la semaine comprendra 5 jours ouvrables de 2 × 8 heures, des heures supplémentaires seront possibles.

La simulation d'un an ou « histoire » sera répétée un certain nombre de fois pour vérifier que les bilans obtenus ne sont pas trop dispersés.

Pour réaliser des échantillons artificiels de toutes les courbes de survie, délais de réparation, et autres valeurs qui sont aléatoires, le tirage peut se faire à partir d'une table de nombres aléatoires équiprobables ou qui sont engendrés directement par le calculateur électronique en sous-programme. On obtient ainsi une suite aléatoire d'intervalles de pannes, ou de délais selon la loi de survie ou de délais considérée (fig. 3.1.).

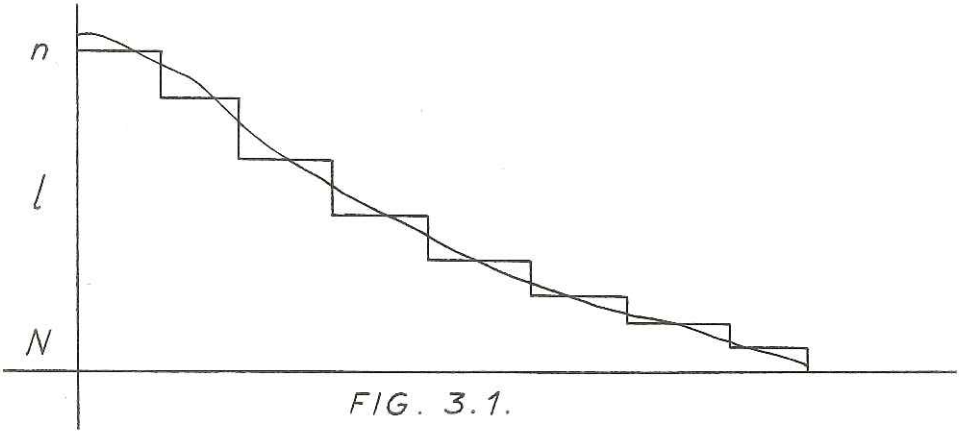

FIG. 3.1.

Les semaines de réparation sont calculées en heures. Quand le total dépasse 40 heures, on répartira le temps d'immobilisation sur la semaine suivante, où il viendra s'accumuler aux heures de cette semaine.

2.3.4. — *Dossier technique.*

Caractéristiques pièces / Machines	Vm^o en h	$\overline{\tau}$ en h	en j^r	$\overline{n}$	pour $t = 1$ an $e^{-\lambda t}$	Λ	λ
A. 1. pièce d'usure1. — ex. bague en bronze	1.000	20	50	2	0,9	—	0,3
2. pièce fatigue — ex. axe	600	2	2.000	3	0,9	—	0,9
3. petites pièces	—	1	5	2	0,85	1/50	—
B. 1. pièce usure (came)	800	4	600	2	0,9	—	0,3
2. pièce fatigue (support moteur)	600	2	110	2	1 ($\alpha = 0$)	—	0,9
3. pièce fatigue	750	13	200	3	1	—	0,9
4. petites pièces		2	10	2	0,8	1/79	—
C. 1. pièce usure (roulement)	2.000	4	400	2	0,95	—	0,3
2. pièce usure (train de pignons)	3.800	8	4.500	3	1	—	0,3
3. pièce fatigue (bielle)	5.400	20	400	3	1	—	0,9
4. petites pièces	—	2	90	2	0,85	1/90	—
D/E. 1. pièce usure (app. électr.)	1.200	18	300	1	0,80	—	0,3
2. petites pièces	—	0,3	5	2	0,85	1/55	—
F. 1. pièce usure (embrayage)	800	12	500	3	0,9	—	0,3
2. pièce fatigue (broche)	1.100	21	10.000	5	1	—	0,9
3. pièce fatigue (contacteur-interrupteur)	400	6	30	1	1	—	0,9
4. pièce fatigue	1.700	0,3	150	1	1	—	0,9
5. petites pièces	—	1	15	2	0,75	1/68	—

La production a lieu en deux équipes de 8 heures soit 16 heures ininterrompues. Quand une machine est en panne, on continue la réparation dans les 8 heures qui suivent, mais il faut doubler le coût horaire d'un ouvrier lorsqu'il travaille en heures supplémentaires.

Les stocks seront évalués toutes les semaines en tenant compte de la production ; celle-ci est évidemment aléatoire à cause des pannes.

2.3.2. — *Notations.*

Nous utilisons les notations suivantes :

λ — indice de dispersion de la loi lognormale.

V_m — durée de vie moyenne pour une loi lognormale.

r — durée de l'indisponibilité à la suite d'une panne.

$\wedge$ — taux de pannes petites pièces par unité de temps (celles-ci suivent une loi exponentielle).

k — taux de production d'une chaîne ou quantité produite par unité de temps.

j_1 — coût de la pièce à remplacer (grandeur fixe).

j_2 — coût moyen de démontage, remontage et réglage (grandeur aléatoire Gaussienne à faible écart type).

n — nombre moyen.

$S'(r)$ — stock à l'entrée de la chaîne à la date (r).

j_3 — coût de stockage d'une matière première, article manufacturé ou produit fini par unité de pièce et unité de temps.

j_4 — coût de pénurie du produit fini par unité de pièce et unité de temps.

α — coefficient caractérisant l'usure générale, tel que $V_m(t) = (1 - \alpha)^n V_m(0)$.
Pour le calcul on admettra la formule limite continue :
$V_m(t) = c^{-\alpha t} V_m(0)$ où α est ajusté de manière qu'il y ait diminution de valeur d'un certain pourcentage pour $t = 1$ an.

Toutes ces variables seront munies d'une indication supplémentaire pour chaque machine.

Par exemple : La durée de vie moyenne pour la pièce d'usure de la machine A : V_m (A).

2.3.3. — *Inventaire des sources de pannes.*

Les machines A, B, C, D, E et F contiennent comme sources de pannes les pièces suivantes :

Machine	Nombre de pièces d'usure	Nombre de pièces de fatigue	Source de pannes dues à de petites pièces
A	1	1	1
B	1	2 distinctes	1
C	2 distinctes	1	1
D	1	0	1
E	1	0	1
F	1	3 distinctes	1

2.3.4. — *Production.*

La chaîne (1) produit la pièce P_1 au taux horaire maximum de $k_1 = 550$; chaque machine de chaîne (2) produit la pièce P_2 au taux horaire maximum de $k_2 = 275$.

La machine F assemble P_1 et P_2 pour former P_3 au taux horaire maximum de 550.

Pour exécuter une pièce P_1 — il faut 48 F. de matière première MP 1
2 F. d'article manufacturé AM 1 (1)
6 F. d'article manufacturé AM 2 (3)
15 F. d'article manufacturé AM 3 (1)

Pour exécuter une pièce P_2 — il faut 25 F. de matière première MP 2
3 F. de matière première MP 3
2 F. d'article manufacturé AM 4
5 F. d'article manufacturé AM 5

Pour exécuter une pièce P_3 — il faut 11 F. de AM 6
6 F. de AM 7

2.3.5. — *Coûts.*

Les coûts proportionnels unitaires en dehors de MP et AM } sont pour $P_1 = 9$ F.
$P_2 = 16$ F.
$P_3 = 21$ F.

Les coûts fixes sont pour la chaîne (1) = 220 F.
(2) = 112 F.
(3) = 115 F.

Le taux de l'argent est de 6 %, pris constant durant toute la période considérée.

Les machines ont un certain âge au temps $t = 0$. Les valeurs sont données dans le tableau suivant.

Machine	Valeur initiale en millions de F.	Age au temps $t = 0$	Valeurs au temps $t = 0$ en millions de F.
A	2,5	2	2
B	2	1	1,8
C	3,5	2	2,8
D	1,5	1	1,35
E	1,5	1	1,35
F	4	3	2,9

Le coût horaire d'un ouvrier chargé de la réparation est de 60 F. (valeur dans laquelle on fait rentrer les coûts fixes). Il est de même quelque soit la réparation, mais est doublé lorsque le travail se fait en heures supplémentaires.

Le nombre des ouvriers charges des réparations est limité à 7. A chaque pièce correspond un nombre moyen n d'ouvriers la distribution des n ouvriers répondant à une distribution binominale.

Les stocks sont évalués à leur coûts de fabrication (valeur de remplacement). Les coûts de stockage sont évalués par l'intérêt du capital investi. Les dépenses de magasinage ne doivent pas intervenir, ces dépenses étant jugées indépendantes du volume. Par contre les stocks doivent intervenir sous forme de contraintes, en spécifiant les limites admissibles.

2.3.6. — *Demande commerciale.*

La demande dans ce cas est considérée comme constante.

2.3.7. — *Fonction économique.*

La fonction économique qui constitue le critère dans la prise de décision entre les diverses politiques possibles est une fonction de coûts à rendre minimum, constituée par la somme des pertes de production — des dépenses de réparation et d'entretien, des amortissements, tenant compte de l'obsolescence aléatoire, des dépenses de stockage. Ces dépenses sont actualisées.

2.3.8. — *Politiques à tester.*

Les diverses politiques sont relatives à l'entretien préventif, la première politique étant celle où l'on ne fait pas d'entretien préventif. L'entretien préventif n'est à considérer que pour les pièces d'usure.

On peut déposer systématiquement à 0,8 $Vm^{(r)}$ où $Vm^{(r)}$ est la durée de vie moyenne au début de l'année (r), alors $\overline{\tau}$ devient $a\overline{\tau}$ avec une certaine valeur de a. On peut également n'appliquer la substitution préventive que pour les pièces ayant un $r\tau > h$ heures.

Pour les machines D et E on peut admettre le remplacement par groupe. Si l'une des pièces de ces machines atteint un âge 0,7 $Vm^{(n)}$ pendant que l'autre a un âge compris entre 0,7 $Vm^{(n)}$ et $Vm^{(n)}$, on remplace les deux pièces, le temps total τ_D et τ_E sont alors réduits de 40 %. On admettra dans ce cas que lors d'un remplacement préventif, les temps τ ne sont plus aléatoires mais ont pour valeur τ.

§ 3. MISE EN MACHINE D'UNE SIMULATION.

Pour résoudre ce problème synthétique de gestion dont la complication voulue est assez grande, il faut utiliser la simulation. Celle-ci est assez coûteuse étant donné le temps de préparation du programme machine ; elle n'a pas été réalisée effectivement sauf les organigrammes. Le modèle synthétique a été construit dans le but de faire apparaître les principales données qui interviennent dans ce genre de problèmes et de comprendre les implications qui les lient.

Une deuxième phase des travaux du « groupe de travail sur l'entretien et la gestion du matériel » consistera en l'étude d'un ou plusieurs cas réels qui seront menés jusqu'à la simulation machine incluse. Un compte-rendu de l'activité du groupe dans cette deuxième phase sera publié ultérieurement.

Méthodes statistiques pour la comparaison de plus de deux traitements I*

Paul De Munter, professeur à l'Université de Tunis.

PLAN

Chapitre 1. — *Comparaison simultanée de* k > 2 *traitements. Hypothèses d'homogénéité contre alternatives générales et particulières.*

1.1. Formulation du problème. Hypothèses et notations.
1.2. Populations normales homoscédastiques.
1.2.1. Homogénéité contre alternatives générales.
1.2.1.1. Observations non progressives.
1.2.1.2. Observations progressives.
1.2.2. Homogénéité contre alternatives particulières.
1.3. Populations normales hétéroscédastiques.
1.4. Populations non normales.
1.4.1. Homogénéité contre alternatives générales.
1.4.2. Homogénéité contre alternatives particulières.
1.5. Références bibliographiques pour ce chapitre.

Remarque : Les autres chapitres paraîtront dans les numéros suivants de cette revue.

Ce sont les :

Chapitre 2 : Sélection du « meilleur » traitement.

Chapitre 3 : Comparaison simultanée de k > 2 traitements à un témoin.

Chapitre 4 : Comparaison de k > 2 traitements deux-à-deux — Classement des traitements en deux groupes — Tests de contrastes — (Populations normales homoscédastiques).

Au cours des dernières années, les méthodes d'analyse statistique d'un groupe de plus de deux moyennes traitements se sont développées et perfectionnées ; des tables numériques nouvelles ont été mises à la disposition de l'expérimentateur.

Nous reprenons la question à partir du test F basé sur l'analyse de la variance dans le cas normal et passons en revue les différents problèmes pratiques qui se posent à l'expérimentateur et pour lesquels une solution existe actuellement (par exemple : comparaison des traitements deux-à-deux, sélection du « meilleur », comparaison à un témoin). Nous citons les conditions d'applicabilité des méthodes et leurs qualités ; nous nous attachons particulièrement au problème de la détermination a priori de l'effectif des observations. Les méthodes sont classées, pour chaque problème pratique, suivant les conditions dans lesquelles elles sont appliquées et suivant le but qu'il faut atteindre (par exemple : cas normal ou non-normal, homogénéité des variances, effectif fixe ou progressif, test d'hypothèse ou intervalle de confiance).

Notre objectif n'est pas d'apporter une contribution originale de nature mathématique, quoique certaines méthodes ou conditions soient nouvelles (leur justification paraîtra ultérieurement). Nous voulons aider l'expérimentateur en lui permettant, une fois le problème posé, de faire l'inventaire des méthodes et de les appliquer en toute connaissance de cause.

Nous remercions Monsieur le Professeur Jean Teghem d'avoir bien voulu lire le manuscrit et d'y avoir introduit de nombreuses améliorations.

(*) Cet article est le premier d'une série qui paraîtra dans les prochains numéros de cette revue.

1. — COMPARAISON SIMULTANEE DE $k > 2$ TRAITEMENTS. HYPOTHESES D'HOMOGENEITE CONTRE ALTERNATIVES GENERALES ET PARTICULIERES.

1.1. *Formulation du problème. Hypothèses et notations.*

a) Alternatives générales et particulières.

Nous savons que la notion de traitement est équivalente à la notion combinée de population — fonction de distribution (De Munter, 1960). Notons π_i la population des réponses possibles au i^e traitement et soit $F_i(x)$ la fonction de distribution de π_i, $(i = 1, 2, \ldots\ldots, k)$. L'égalité des traitements est donc un fait équivalent à l'égalité des fonctions F_i. D'une manière générale, on est amené à tester l'hypothèse

$$H_o : F_1(x) = F_2(x) = \ldots F_k(x),$$

pour tout x. L'hypothèse H_o exprime l'homogénéité des populations π_i. Lorsque H_o est vraie, il n'y a plus k populations, mais une seule, dont la fonction de distribution est la fonction $F_i(x)$ par exemple.

Lorsqu'une méthode statistique quelconque conduit à rejeter H_o, on est en droit de penser qu'une alternative H à H_o est vraie. Or, si l'on admet qu'une alternative H exprime seulement qu'une au moins des égalités entre les F_i tombe en défaut, il y a $2^k - 1$ alternatives H. S'il y a 5 traitements en présence, il y a 31 alternatives à l'hypothèse d'homogénéité. Nous dirons des alternatives H qu'elles sont générales.

Parmi les alternatives générales, il peut y en avoir que l'expérimentateur sait, a priori, être impossibles. Il peut en être de plus ou moins vraisemblables. Enfin, il se peut que l'expérimentateur puisse, a priori, déterminer un ensemble restreint d'alternatives seules possibles. Par exemple, seules peuvent être possibles les alternatives H : $F_1 < F_2 < \ldots < F_k$. Nous dirons dans ce cas que les alternatives sont particulières.

Il est (intuitivement) justifié de croire qu'un test statistique de H_o contre les alternatives générales est moins efficace qu'un test de H_o contre des alternatives particulières. Un test de H_o contre H_g (alternatives générales) rejetera H_o moins facilement qu'un test de H_o contre H_p (alternatives particulières) surtout si, justement, c'est H_p qui est vraie.

Notons D_o la décision statistique de considérer H_o comme l'hypothèse vraie et soit D la décision statistique de considérer qu'une alternative H (générale ou particulière) est vraie. Prendre la décision D_o revient à accepter H_o et prendre la décision D revient à rejeter H_o en faveur d'une alternative H générale ou particulière.

b) Détermination a priori de l'effectif échantillon.

D'une manière générale, on fixe a priori la probabilité de prendre la décision D_o lorsque H_o est vraie, soit $P\{D_o|H_o\}$. Nous appellerons « condition (C_2) » la nécessité pour $P\{D_o|H_o\}$ d'avoir une certaine valeur fixée à l'avance ; on aura généralement

$$(C_2) \equiv P\{ D_o \mid H_o \} = 1 - \alpha \qquad (= 0{,}95),$$

où α s'appelle « niveau de signification du test » ou « erreur de première espèce ». Mais il est tout aussi important de fixer a priori la probabilité de prendre la décision D lorsque H est vraie, soit $P\{D|H\}$. Nous appellerons « condition (C_1) » la nécessité pour $P\{D|H\}$ d'avoir une valeur fixée à l'avance ; on aura généralement

$$(C_1) \equiv P\{ D \mid H \} = 1 - \beta \qquad (= 0{,}95 \text{ ou } 0{,}90),$$

où β s'appelle « erreur de deuxième espèce ».

Les symboles α et β ont effectivement la signification d'une erreur. En effet, on a

$$\alpha = 1 - P\{ D_o \mid H_o \} = P\{ D \mid H_o \},$$

c'est-à-dire, la probabilité de rejeter H_o alors que H_o est l'hypothèse réellement vraie. D'autre part, on a

$$\beta = 1 - P\{ D \mid H \} = P\{ D_o \mid H \},$$

c'est-à-dire, la probabilité d'accepter H_o alors qu'une alternative H est vraie.

Pour exploiter la condition (C_1), il faut préciser H. Par exemple, H : $F_1(x) = F_2(x) = \ldots = F_{k-1}(x) = F(x)$, $F_k(x) = F(x - c)$, exprimant que les $k - 1$ premiers traitements sont identiques et que le k^e diffère, seulement en ce qui concerne sa moyenne, d'une quantité c bien précisée. Pour exploiter la condition (C_1) il faut donc particulariser H. Il est évident que cette praticularisation se fait dans le sens le plus plausible. Si, par exemple, on doit comparer trois engrais dont l'un diffère, par ses qualités chimiques, nettement des deux autres, les alternatives à H_0 sont du type cité plus haut. Dans ces conditions, on peut tirer de la condition (C_1) le nombre d'observations qu'il convient de prélever dans chaque population. (C'est la raison pour laquelle nous avons affecté à la condition sur l'erreur de deuxième espèce l'indice « 1 » : on commence par fixer α et β d'où l'on tire tout d'abord l'effectif échantillon afin de pouvoir entreprendre l'essai. Ensuite, de la condition (C_2) on tire la règle du test qui permet de prendre une décision). Mais il ne faut pas perdre de vue que cette manière de faire introduit une nouvelle sorte d'erreur. Il a fallu, pour calculer $P\{D|H\}$ faire un choix parmi les alternatives H. Soit H' l'alternative choisie et D' la décision de considérer H' comme vraie. On a donc $P\{D'|H'\} = 1 - \beta$. Mais cela ne donne aucune garantie de prendre la décision D' si H''($\neq$ H') est vraie.

c) Méthodes progressives.

Il est possible de fixer a priori $P\{D_0|H_0\}$ et $P\{D|H\}$ sans pour autant que cela impose un effectif échantillon fixe. Alors, la méthode est dite progressive (ou séquentielle). Cela signifie que l'effectif échantillon est aléatoire ; il dépend du comportement de la variable observée. Les observations sont prélevées successivement et à chaque prélèvement on est amené à prendre l'une des trois décisions : D_0, D, faire un nouveau prélèvement. Les méthodes progressives pour tester H_0 contre H ne sont ni nombreuses, ni développées. Nous en décrivons une due à Johnson (1953) et améliorée par Ray (1956).

d) Homogénéité relative.

Supposons que l'on soit amener à comparer trois engrais dont l'application a pour seul but d'accroître le rendement. Et l'on suppose que ce but est atteint et seulement celui-là. En d'autres termes, l'application de ces engrais ne perturbe pas, par exemple, la dispersion des réponses. Dans ces conditions, la moyenne de la population des réponses à l'un des engrais en est la caractéristique essentielle.

Mais il se peut que les populations des réponses possibles aux différents engrais aient des fonctions de distribution de formes différentes. Cela n'intéresse pas l'expérimentateur. Ce qui l'intéresse essentiellement, c'est le comportement respectif des moyennes de ces populations. Dans ce cas, l'homogénéité des traitements est équivalente à une homogénéité relative des fonctions de distribution, relative aux valeurs moyennes seulement. L'hypothèse généralement testée est $H_0 : \mu_1 = \mu_2 = \ldots = \mu_k$, où μ_i est la moyenne population du i^e traitement.

e) Hypothèse de normalité.

Dans ce qui précède, rien n'a été dit ni supposé quant à la forme des fonctions de distribution $F_i(x)$ des populations π_i.

Lorsqu'il est concevable (*i*) que la variable aléatoire observée puisse prendre n'importe quelle valeur réelle, (*ii*) que la variable aléatoire observée résulte de l'addition d'un très grand nombre d'impulsions élémentaires représentatives de facteurs aléatoires plus ou moins indépendants, il est justifié de supposer que la fonction de distribution de cette variable est normale ou approximativement normale. Rappelons que si X est une variable normale,

$$P\{X \leqslant x\} = \frac{1}{\sqrt{2\pi}\,\sigma} \int_{-\infty}^{x} e^{-\frac{1}{2}\left(\frac{t-\mu}{\sigma}\right)^2} dt ,$$

où μ est la moyenne et σ l'écart-type de X ; on note X est $N(\mu,\sigma)$.

Considérons les k traitements et les populations π_i des réponses possibles. S'il est concevable d'admettre que les réponses possibles statisfont aux conditions (*i*) et (*ii*), il est justifié de supposer que la fonction de distribution de π_i est normale de moyenne μ_i et d'écart-type σ_i. Si les populations ont toutes la même variance ($\sigma_1^2 = \sigma_2^2 = \ldots \sigma_k^2$), nous dirons qu'elles sont homoscédastiques ; sinon elles sont hétéroscédastiques.

Nous envisagerons tout d'abord le cas de l'homoscédasticité. Il est fort important d'en bien comprendre la portée sur la nature des traitements. Exemple : Une usine de produits pharmaceutiques met au point trois fortifiants A, B et C pour enfants. Elle désire ne mettre sur le marché que le « meilleur ». Les traitements sont appliqués à trois groupes d'enfants (20 enfants par groupe, par exemple) pendant un mois. Ces enfants ont été choisi au hasard dans une population dont tous les éléments satisfont à certains critères portant sur l'âge, le poids initial, le standing social, l'état de santé, etc. De telle sorte, qu'on a éliminé les facteurs principaux d'hétérogénéité systématique pouvant affecter les réponses aux traitements. Dès l'application des traitements, la population unique des enfants se scinde en trois, respectivement caractéristique de chaque traitement. Supposer que ces populations sont normales implique une hypothèse sur les traitements : les réponses se répartissent symétriquement autour de la moyenne ; il y a autant de petits accroissements de poids que de grands. Supposer en outre que les fonctions de distribution sont homoscédastiques implique que les réponses sont également dispersées. Or, le meilleur traitement est également celui qui agit le plus indépendamment des individus, c'est-à-dire, celui qui a la dispersion la plus faible. Donc faire l'hypothèse de l'homoscédasticité, c'est faire a priori une hypothèse sur la qualité des traitements (cet exemple est repris au chapitre 2, § 1).

Admettons qu'il soit permis de faire l'hypothèse que les fonctions de distribution des populations des réponses possibles aux différents traitements soient normales homoscédastiques (ce qui se note « π_i est $N(\mu_i, \sigma)$ », μ_i étant la moyenne et σ l'écart-type commun). Les populations ne diffèrent que par les valeurs moyennes μ_i. L'hypothèse statistique $H_o : \mu_1 = \mu_2 = \ldots = \mu_k$ est donc équivalente à l'hypothèse d'homogénéité.

Dans le cas de l'hétéroscédasticité, les fonctions de distribution des populations des réponses possibles sont $N(\mu_i, \sigma_i)$ et l'hypothèse $H_o : \mu_1 = \mu_2 = \ldots = \mu_k$ n'implique qu'une homogénéité relative des fonctions de distribution.

Les alternatives à $H_o : \mu_1 = \mu_2 = \ldots = \mu_k$ sont soit générales et impliquent qu'un signe égal au moins tombe en défaut, soit particulières et sont, par exemple, $H_p : \mu_1 \geqslant \mu_2 \geqslant \ldots \geqslant \mu_k$ où un signe égal au moins n'existe pas.

Il est souvent possible, par une transformation adéquate des observations, de rendre homoscédastiques des fonctions de distribution présentant une certaine hétérogénéité de variance (De Munter, 1958).

f) Non-normalité.

Lorsqu'il n'est pas permis de supposer que les fonctions de distribution des populations sont normales, il peut également être possible, par une transformation adéquate des données, de les normaliser (De Munter, 1958). Sinon, on applique des méthodes valables pour n'importe quelles fonctions de distribution. Dans ce cas, on teste l'hypothèse d'homogénéité $H_o : F_1 = F_2 = \ldots = F_k$ contre H_g ou H_p.

1.2. *Populations normales homoscédastiques.*

1.2.1. *Hypothèse d'homogénéité contre alternatives générales.*

1.2.1.1. *Observations non-progressives.*

Soit π_i la population des réponses possibles au i^e traitement, ($i = 1, 2, \ldots, k$). La fonction de distribution de π_i est $N(\mu_i, \sigma)$. Soit Y_{ij} la réponse de la j^e unité expérimentale ayant subi le i^e traitement. On suppose que Y_{ij} est la somme d'une moyenne générale μ, d'une impulsion τ_i due essentiellement au i^e traitement et d'une variable aléatoire X_{ij} dont la fonction de distribution est $N(O, \sigma)$. On a donc

$$Y_{ij} = \mu + \tau_i + X_{ij} ,$$

où X_{ij} résulte de l'influence de tous les facteurs aléatoires de perturbation autres que le i^e traitement. On appelle donc X_{ij} l'erreur affectant la réponse de la j^e unité expérimen-

tale ayant subi le i^e traitement. Les erreurs X_{ij} sont des variables normales homoscédastiques indépendantes. On peut généralement faire cette hypothèse d'indépendance grâce à une répartition aléatoire des traitements parmi les unités expérimentales : les facteurs non-contrôlés agissent alors sans tendance privilégiée. Le terme μ a la signification d'une moyenne générale perce que, d'une part les τ_i sont tels que $\Sigma \tau_i = 0$, et d'autre part X_{ij} a pour moyenne zéro. Il en résulte que la moyenne du i^e traitement est

$$\mu_i = \mu + \tau_i$$

et que $\Sigma \mu_i / k = \mu$; par conséquent, la différence entre les moyennes des traitements (1) et (2), par exemple, est égale à $\tau_1 - \tau_2$.

L'égalité des moyennes traitements est réalisée lorsque l'hypothèse H_o : $\tau_1 = \tau_2 = \ldots = \tau_k = 0$ est vraie. L'hypothèse H_o exprime donc l'homogénéité des populations normales. Une alternative H_g est générale si elle exprime seulement que, pour deux valeurs de i au moins, $\tau_i \neq 0$. Tout écart entre H_o et H_g peut donc se mesurer par le coefficient

$$\lambda = \sqrt{\frac{\Sigma \tau_j^2}{k \sigma^2}}$$

Si par exemple, $\tau_1 = \tau_2 = \ldots = \tau_{k-2} = 0 \quad \tau_{k-1}$ et $\tau_k \neq 0$, on a

$$\lambda = \sqrt{\frac{\tau_{k-1}^2 + \tau_{k-2}^2}{k \sigma^2}}$$

L'estimation de λ, a priori, est un élément important de l'analyse statistique. Dans un essai variétal, par exemple, l'expérimentateur doit comparer trois variétés A, B et C de blé du point de vue du rendement en quantité de grain. Si rien n'a jamais été expérimenté sur le sujet, si l'expérimentateur n'a absolument aucune information a priori, il doit organiser un essai d'orientation qui fournira des estimations de σ et des moyennes μ_i. Mais généralement, l'expérimentateur possède une certaine information et par conséquent une estimation de σ. Supposons, pour simplifier l'exposé, que l'expérimentateur sache à l'avance que deux variétés, A et B par exemple, sont presque identiques, tandis que la troisième se distingue des deux autres. Une certaine différence minimum entre C et (A , B) est importante (du point de vue économique, par exemple). En d'autres termes, si le rendement moyen de C dépasse les rendements moyens de A et B d'une quantité $\Delta \geqslant \Delta^*$, C présente un avantage économique certain.

On teste donc H_o : $\tau_A = \tau_B = \tau_C$ contre H_g : $\tau_A = \tau_B = \tau \quad \tau_C = \tau + \Delta^*$, et il convient de prendre, avec une probabilité élevée, la décision D (accepter H_g) si H_g est vraie. Pour satisfaire à cette condition (C_1), (voir § 1.1.b), il importe de calculer λ. Sachant qu'il faut que $\tau_A + \tau_B + \tau_C = 0$, d'où l'on a $\Delta^* = -3\tau$, il vient

$$\lambda = \frac{\sqrt{2}}{3} \frac{\Delta^*}{\sigma}$$

On constate que la détermination de λ nécessite une information d'ordre statistique (σ ou une estimation de σ) et une information inhérente au problème essentiel que l'on veut résoudre (une valeur de Δ^* conditionnée par des impératifs économiques, techniques, etc.).

Dans l'exposé de la méthode statistique pour tester H_o contre H_g , nous nous limitons tout d'abord au cas où il y a le même nombre n d'observations prélevées dans chaque population. Il y a donc $kn = n'$ observations au total. Soit également $Y_{i.}$. La somme des observations relatives à la population π_i et soit $\overline{Y}_{i.} = Y_{i.}/n$ la moyenne. Soit enfin Y.. la somme totale de toutes les observations et $\overline{Y}.. = Y../n'$ la moyenne générale. Notons $C = Y^2../n'$ un terme correctif qui interviendra dans la suite.

La variation totale de l'essai est

$$S_T = \sum_{ij} (Y_{ij} - Y..)^2 = \sum_{ij} Y^2_{ij} - C .$$

La variation (totale) due aux traitements est

$$S_t = n \sum_i (\overline{Y}i. - \overline{Y}..)^2 = \frac{1}{n} \sum_i Y^2_{i.} - C$$

tandis que la variation (totale) due aux erreurs est

$$S_e = \sum_{ij} (Y_{ij} - \overline{Y}_{i.})^2 = S_T - S_t = \sum_{ij} Y^2_{ij} - \frac{1}{n} \sum_i Y^2_{i.}$$

Les variations moyennes correspondantes sont

$$V^2_t = S_t/(k-1) \qquad \text{et} \qquad V^2_e = S_e/(n'-k)\,,$$

dont le rapport $F = V^2_t/V^2_e$ est la variable de Fisher-Snedecor avec $(k-1)$ et $(n'-k)$ degrés de liberté.

On prend la décision D_o si la valeur observée f de la variable aléatoire F est telle que (°)

$$f < f_\alpha\ ;$$

on prend la décision D si $f > f_\alpha$. Le nombre f_α est tiré des tables de la distribution de Fisher-Snedecor ; il est tel que la condition (C_2) soit satisfaite (voir plus loin et § 1.1.b).

L'effectif n commun à tous les échantillons est déterminé a priori de telle sorte que la probabilité de prendre la décision D alors que H_g est vraie soit fixée à l'avance, à savoir,

$$P\{\, D \mid H_g \,\} = 1 - \beta\,, \qquad (C_1)$$

où H_g est spécifiée par une valeur de λ. Nous reproduisons aux tableaux ci-contre les valeurs de n, correspondant à différents λ, telles que la condition (C_1) soit satisfaite lorsque $\alpha = 0{,}05$ et $\beta = 0{,}10$ (table 1.1.), $\beta = 0{,}05$ (table 1.2).

k \ λ	1	0,9	0,8	0,7	0,6	0,5	0,4	0,3
3	6	7	8	10	12	18	27	47
4	5	6	7	9	11	15	23	40
5	5	5	6	7	10	11	21	35

Table 1.1. Valeurs de n correspondant à différentes valeurs de λ et de k, telles que $P\{\, D_o \mid H_o \,\} = 0{,}95$ et $P\{\, D \mid H_g \,\} = 0{,}90$.

k \ λ	1	0,9	0,8	0,7	0,6	0,5	0,4	0,3
3	7	8	9	11	15	21	34	57
4	6	7	8	10	13	17	27	48
5	6	6	7	9	12	16	24	43

Table 1.2. Valeurs de n correspondant à différentes valeurs de λ et de k, telles que $P\{\, D_o \mid H_o \,\} = 0{,}95$ et $P\{\, D \mid H_g \,\} = 0{,}95$.

(°) Les majuscules sont des variables aléatoires dons les réalisations sont représentées par les minuscules correspondantes.

Les valeurs de n fournies par les tables 1.1. et 1.2 ont été tirées des abaques de Feldt et Mahmoud (1958).

La valeur f_α, fournie par les tables de la distribution de Fisher-Snedecor avec $(k-1)$ et $(n'-k)$ degrés de liberté, est telle que

$$P\{ D_o \mid H_o \} = 0{,}95 . \qquad (C_2)$$

Etant donné une valeur observée f de F, il est souvent utile de déterminer $P\{ F > f \}$. Il est évident qu'on prend la décision D si $P\{ > f \} < \alpha$ (généralement 0,05). Les tables de la distribution de Fisher-Snedecor fournissent les f_α tel que $P\{ F > f_\alpha \} = \alpha$ mais ne fournissent pas la possibilité de déterminer $P\{ F > f \}$ quel que soit f. A cette fin, il faut disposer des tables de la fonction bêta incomplète (Pearson, 1956). Ces tables fournissent les valeurs de $I_x(p, q)$ pour différentes valeurs de x, p et q. Supposons que la variable F ait une distribution de Fisher-Snedecor avec f_1 et f_2 degrés de liberté. On a

$$P\{ F > f \} = I_x\left(\frac{f_2}{2}, \frac{f_1}{2} \right)$$

où $x = f_2/(f_2 + f_1 f)$.

Le premier exemple numérique est relatif à des données artificielles. Soient trois traitements A, B et C dont les réponses possibles constituent des populations normales homoscédastiques. Afin que l'homoscédasticité ne soit pas mise en doute, les variances échantillons s_i^2 sont toutes égales. Les données sont rassemblées à la table 1.3. Les effectifs échantillons ($n = 6$) ont été fixés sans considération de la condition (C_1). Il n'empêche qu'il convient de calculer $P\{ D \mid H_g \}$. Les moyennes populations sont $\mu_1 = \mu + \tau_1$, $\mu_2 = \mu + \tau_2$ et $\mu_3 = \mu + \tau_3$ où μ est estimée par $\bar{y}_{..} = 12$. Posons $P\{ D_o \mid H_o \} = 0{,}95$; quelle est, dans ces conditions, $P\{ D \mid H_g \}$ si H_g : $\tau_1 = -1$, $\tau_2 = -1$ et $\tau_3 = 2$? La valeur de λ qui correspond à H_g est $\sqrt{2/\sigma^2}$, où l'on peut remplacer σ^2 par la variance échantillon, $s_i^2 = 2$; donc $\lambda = 1$. Par conséquent, en vertu de la table 1.1., on a $P\{ D \mid H_g \} = 0{,}90$.

	y_{ij}						n	$y_{i.}$	$\bar{y}_{i.}$	s_i^2
A	10	10	12	8	9	11	6	60	10	2
B	11	9	13	11	10	12	6	66	11	2
C	15	15	13	17	14	16	6	90	15	2
							18	216	12	

Table 1.3. Données du premier exemple.

Le terme correctif (°) a pour valeur $c = (216)^2/18 = 2592$. La variation totale a pour valeur $s_T = (10)^2 + (10)^2 + \ldots + (14)^2 + (16)^2 - 2592 = 114$. La variation (totale) due aux traitements a pour valeur $s_t = (90)^2 - 2592 = 84$. Par conséquent, la variation (totale) due aux erreurs vaut $s_e = s_T - s_t = 30$. Les variations moyennes correspondantes sont $v_t = s_t/2 = 42$ et $v_e^2 = s_e/15 = 2$. On remarque qu'effectivement, la variation moyenne résiduelle (estimation correcte de σ^2) est égale à la variance (commune) correcte échantillon. La réalisation de F vaut $f = v_t^2/v_e^2 = 21$. La distribution de la variable F avec 2 et 15 degrés de liberté fournit la valeur $f_\alpha = 3{,}68$ pour $\alpha = 0{,}05$, c'est-

(°) La minuscule c est la réalisation du terme correctif $C = Y_{..}^2/n'$; de même pour $s_T, s_t, \ldots$, réalisations de $S_T, S_t, \ldots$

à-dire, pour que $P\{D_o|H_o\} = 0{,}95$. Par conséquent, on est amené à prendre la décision D : rejeter H_o en faveur de H_g. Les tables de distribution de F ne permettent pas de calculer $P\{F > 21\}$. Mais on a, d'après les tables de la fonction bêta incomplète (Pearson, 1956), $P\{F > 21\} = I_x(15/2, 1)$, où $x = 15/57 = 0{,}263157$ et par conséquent, $P\{F > 21\} = 45 \cdot 10^{-6}$.

Considérons un second exemple numérique étroitement lié au précédent : les données relatives au traitement C sont diminuées d'une quantité fixe égale à 2,84. Cette fois $y_3 = 72{,}96$ (au lieu de 90) et $\bar{y}_3 = 12{,}16$ (au lieu de 15) ; mais il est évident que s_3^2 garde la même valeur 2. On calcule successivement $c = 2199{,}17$, $s_T = 44{,}02$, $s_t = 14{,}02$, $s_e = 30$, $v_t^2 = 7{,}01$, $v_e^2 = 2$ et finalement $f = 3{,}5$. Cette fois, la valeur observée pour $F\,(f = 3{,}5)$ est inférieure au seuil 5 % (3,68). On prend donc la décision D_o.

Certains auteurs ont mis en évidence (voir, par exemple, David & Johnson, 1951 et Hack, 1958) que le test-F est « robuste » contre la non-normalité des populations étudiées. Autrement dit, $P\{D_o|H_o\}$ n'est pas éloigné de 0,95 (lorsque f_z correspond à $\alpha = 0{,}05$) dans certains cas de non-normalité même très accentuée ; d'autre part, $P\{D|H_g\}$ n'est également pas fort influencé par la non-normalité. On a constaté également (voir, par exemple, Horsnell, 1953 et Box, 1953) que l'hétéroscédasticité des populations n'a d'importance pratique que dans le cas où les effectifs échantillons sont différents (ce problème est repris au § 1.3.).

Lorsque les effectifs échantillons sont différents, les tables 1.1 et 1.2 ne peuvent être utilisées et par conséquent, il n'est pas possible d'exploiter la condition (C_t). L'analyse statistique n'est modifiée qu'en ce qui concerne la définition de S_t ; on a

$$S_t = \sum_i n_i\,(\overline{Y}_{i.} - \overline{Y}_{..})^2 = \sum \frac{Y_{i.}^2}{n_i} - C$$

où n_i est l'effectif de l'échantillon prélevé dans la population π_i.

1.2.1.2. *Observations progressives.*

Les hypothèses de travail et les notations sont celles introduites au § 1.2.1.1. Rappelons ici que la réponse de la j^e unité expérimentale au i^e traitement est $Y_{ij} = \mu + \tau_i + X_{ij}$, où $\Sigma\,\tau_i = 0$ et où X_{ij} est une variable $N(o, \sigma)$. L'hypothèse d'homogénéité est $H_o : \tau_1 = \tau_2 = \ldots = \tau_k = 0$, les alternatives générales étant spécifiées par les valeurs du paramètre $\lambda = \sqrt{\Sigma\,\tau_i^2/k\,\sigma^2}$. D_o est la décision de considérer H_o vrai ; D est la décision de considérer H_g vrai. Dans le cas de l'expérimentation non-progressive, on fixe l'effectif commun n de chaque échantillon a priori de telle sorte que si $P\{D_o|H_o\} = 0{,}95$ on ait soit $P\{D|H_g\} = 0{,}90$, soit $P\{D|H_g\} = 0{,}95$.

L'expérimentation progressive consiste en l'observation successive des unités expérimentales. Dans le cas du test que nous exposons ici, le démarrage se fait en observant simultanément 3 unités si $k = 3{,}5$ ou 7 et 4 unités si $k = 4{,}6$ ou 8, dans chaque population. Après cette première étape, on prend soit la décision D_o, soit la décision D, soit la décision D_e de continuer l'expérimentation. Dans ce dernier cas, on fait deux observations simultanées dans chaque population ; c'est la deuxième étape qui se solde par D_o, D ou D_e, etc. (Le fait de devoir faire deux observations à chaque étape est imposé seulement par la nature des tables numériques existantes ; rien n'empêche d'interpoler (graphiquement) dans les tables et de ne faire qu'une seule observation à chaque étape).

L'effectif commun échantillon, à chaque étape, est une variable aléatoire que nous notons N ; l'effectif total correspondant est $N' = kN$.

A chaque étape, on calcule $S_T^{(N)}$ et $S_t^{(N)}$, les variations totale et due aux traitements, correspondant à N. Le calcul de ces variations est en tout point analogue à celui exposé dans le cas où n est fixe (§1.2.1.1.). Ensuite, on calcule

$$G^{(N)} = S_t^{(N)} / \left(S_T^{(N)} - S_t^{(N)}\right)$$

Soit $g^{(n)}$ la réalisation de $G^{(N)}$ sur l'essai, au moment où $N = n$. Remarquons que

$$g^{(n)} = \frac{k-1}{n'-k} f\,,$$

λ	n	$k = 4$		$k = 6$		n	$k = 3$		$k = 5$		$k = 7$	
		$10^3\,a$	$10^3\,b$	$10^3\,a$	$10^3\,b$		$10^3\,a$	$10^3\,b$	$10^3\,a$	$10^3\,b$	$10^3\,a$	$10^3\,b$
0,707	4	65	1825	168	1272	3			72	4176	184	2407
	6	110	770	183	646	5	37	1319	142	927	211	784
	8	126	521	183	464	7	72	696	155	568	207	512
	10	131	411	180	377	9	89	497	158	432	199	401
	12	134	350	176	327	11	99	400	159	360		
	14	136	310			13	105	344	159	317		
	16	138	382			15	108	306	157	287		
1	4	266	1498	398	1237	3	124	4760	331	2469	469	1925
	6	284	898	373	763	5	195	1174	340	973	420	888
	8	287	637	354	601	7	221	762	331	687	387	645
	10	286	540	340	521	9	234	605	322	565	365	552
	12	284	482	330	473	11	240	522	314	498	349	494
	14			322	340	13	244	471	308	456	338	456
	16	280	424			15	247	436	303	428	328	430
1,414	4	616	1696	765	1535	3	458	3361	747	2445		
	6	596	1128	685	1081	5	497	1403	677	1272		
	8	583	935	645	913	7	507	1040	637	994		
	10	574	837	620	825	9	510	891	613	869		
	12	567	777	602	770	11	512	809	596	799		
	14	557	708	578	706	13	512	758	584	753		
	16					15	511	724	575	720		

Table 1.4. — Nombres a et b tels que $P\{D_o|H_o\} = 0{,}95$ et $P\{D|H_g\} = 0{,}95$.

où f est la réalisation correspondante de F. On prend la décision

$$D_o \quad \text{si} \quad g^{(n)} \leqslant a \ ;$$

$$D \quad \text{si} \quad g^{(n)} \geqslant b \ ;$$

$$D_c \quad \text{si} \quad a < g^{(n)} < b \ ;$$

les nombres a et b faisant l'objet de la table 1.4. (Cette table est un abrégé des résultats numériques de Ray, 1956 ; la méthode elle-même vient de Johnson, 1953). Les seuils a et b sont tels que $P\{D_c|H_o\} = 0{,}95$ et $P\{D|H_g\} = 0{,}95$.

L'effectif commun N des échantillons étant une variable aléatoire, on peut, théoriquement tout au moins, en déterminer la valeur moyenne. Des résultats numériques de Ray (1956), il résulte que la valeur moyenne de N est inférieure à l'effectif fixe correspondant, d'une quantité égale, grosso modo, à 35 % de cet effectif fixe. Par exemple, dans le cas où il y a trois traitements ($k = 3$) et où $P\{D_c|H_o\} = 0{,}95$ et $P\{D|\lambda = 1\} = 0{,}95$, il faut fixer a priori le nombre d'observations à 7 (voir table 1.2 du § 1.2.1.1.) ; dans les mêmes conditions, la valeur moyenne de N vaut approximativement 4.

Appliquons la méthode au premier exemple du § 1.2.1.1, en supposant que les observations (table 1.3) sont dans l'ordre de leur prélèvement successif. La première étape est le calcul de $g^{(3)}$ pour les trois premières données du tableau 1.3 que nous avons reproduites ci-contre. (On démarre avec 3 observations, parce que $k = 3$ est impair).

A	10	10	12	3	32
B	11	9	13	3	33
C	15	15	13	3	43
				9	108

Il vient successivement $c^{(3)} = \dfrac{(108)^2}{9} = 1296$;

$$s_T^{(3)} = (10)^2 + (10)^2 + \ldots + (13)^2 - 1296 = 38 \ ;$$

$$s_t^{(3)} = \frac{1}{3}[(32)^2 + (33)^2 + (43)^2] - 1296 = 24{,}67 \ ;$$

et finalement, $g^{(3)} = s_t^{(3)}/(s_T^{(3)} - s_t^{(3)}) = 1{,}850$. Or, pour $k = 3$ et $n = 3$ (et $\lambda = 1$ comme dans l'exemple original), $a = 0{,}124$ et $b = 4{,}760$. On prend donc la décision D_c et deux nouvelles observations dans chaque population. Mais le prix des observations étant très élevé, il y a intérêt à ne prendre qu'une seule observation dans chaque population. Il faut alors déterminer les seuils a et b par interpolation graphique ; nous verrons d'ailleurs, que seule l'interpolation pour b est indispensable. La deuxième étape est donc le calcul de $g^{(4)}$ pour les données ci-contre.

A	10	10	12	8	4	40
B	11	9	13	11	4	44
C	15	15	13	17	4	60
					12	144

Il vient successivement $c^{(4)} = -\frac{(144)^2}{12} = 1728$;

$$s_T^{(4)} = s_T^{(3)} + (8)^2 + (11)^2 + (17)^2 - 1728 = 80 :$$

$$s_t^{(4)} = \frac{1}{4}[(40)^2 + (44)^2 + (60)^2] - 1728 = 56 ;$$

et finalement, $g^{(4)} = s_t^{(4)}/S_T^{(4)} - S_t^{(4)}) = 2{,}333$. Il est évident qu'il suffit d'interpoler pour b ; on obtient $b \cong 1.800$. Et par conséquent, on prend la décision D. On arrive au même résultat qu'au § 1.2.1.1, mais avec 4 observations seulement, au lieu de 7. (En fait, il n'y en avait que 6, parce que nous avions $P\{D|H_g\} = 0{,}90$ au lieu de 0,95).

1.2.2. *Hypothèse d'homogénéité contre alternatives particulières.*

L'énoncé du problème est identique à celui qui fait l'objet du § 1.2.1 à cela près que cette fois, l'objectif est de tester $H_o : \tau_1 = \tau_2 = \ldots = \tau_k = 0$ contre les alternatives particulières $H_p : \tau_1 \geqslant \tau_2 \geqslant \ldots \geqslant \tau_k$ où au moins une égalité tombe en défaut.

Soient $\overline{y_i}$, $(i = 1, 2, \ldots, k)$, les moyennes calculées des différents échantillons, la moyenne $\overline{y_i}$ correspondant à τ_i. On ordonne les moyennes y_i par valeurs décroissantes (c'est-à-dire, dans l'ordre des τ_i correspondants lorsque H_p est vrai). Chaque fois que deux moyennes consécutives $\overline{y_i}$, $\overline{y_{i+1}}$ sont dans l'ordre inverse à celui de H_p (c'est-à-dire $\overline{y_i} < \overline{y_{i+1}}$) les deux échantillons correspondants sont réunis en un seul échantillon. L'opération de réunion des échantillons contigus est réalisée jusqu'au moment où les moyennes échantillons sont dans l'ordre de H_p. Considérons un exemple (Bartholomew, 1959, I) dont les données sont reproduites au tableau ci-contre.

k	1	2	3	4	5
y_i	110	123	126	140	96
n_i	4	4	4	4	4
$\overline{y_i}$	27,5	30,75	31,5	35	24
				29,5	
			30,17		
		30,31			

L'hypothèse H_p exprime que les moyennes population sont dans l'ordre $\mu_1 \leqslant \mu_2 \leqslant \mu_3 \leqslant \mu_4 \leqslant \mu_5$. Les moyennes échantillons sont dans cet ordre, sauf en ce qui concerne $\overline{y_4}$ et $\overline{y_5}$. Les échantillons (5) et (4) sont d'abord réunis ; l'échantillon global a pour moyenne 29,5 inférieure à la moyenne de l'échantillon (3). On réunit donc (3) à (5 + 4) ; ce nouvel échantillon a pour moyenne 30,17 inférieure à la moyenne de l'échantillon (2). On réunit donc (2) à (5 + 4 + 3). Cette fois, la moyenne du nouvel échantillon (30,31) est supérieure à la moyenne de l'échantillon (1). Par conséquent, les 5 échantillons d'origine sont ramenés à 2.

Soit $l \leqslant k$, le nombre d'échantillons qui subsistent après les opérations successives de réunion. Si les moyennes originales $\overline{y_i}$ sont dans l'ordre de H_p $l = k$. En réalité, le nombre d'échantillons qui subsistent après les opérations de réunion est une variable aléatoire : il dépend de l'ordre des variables aléatoires $\overline{X_i}$. Nous devons donc noter ce

nombre L. Soit d'ailleurs $P_k\{L = l\}$ la probabilité pour que, H_o étant vraie, les k échantillons originaux se réduisent à un groupe de l échantillons. On a évidemment $P_k\{L = 1\} = 1/k$ et $P_k\{L = k\} = 1/k!$ On démontre que, (Miles, 1959),

$$P_k\{L = l\} = [P_{k-1}\{L = l-1\} + (k-1)P_{k-1}\{L = l\}]/k ,$$

$$P_k\{L = l\} = \frac{1}{l}\sum_{j=l-1}^{k-1} P_{k-j}\{L = 1\}P_j\{L = l-1\} .$$

Nous empruntons à Miles (1959), un extrait d'une table fournissant les valeurs de $P_k\{L = l\}$ pour différentes valeurs de l et de k table 1.5).

l	$k =$ 2	3	4	5	6	7	8	9
1	5000	3333	2500	2000	1667	1429	1250	1111
2	5000	5000	4383	4167	3805	3500	3241	3019
3		1667	2500	2917	3125	3222	3257	3255
4			417	833	1181	1458	1679	1854
5				83	208	347	486	619
6					14	42	80	125
7						2	7	15
8							0,2	0,9

Table 1.5. $10^4 . P_k\{L = l\}$.

Supposons donc que les k échantillons d'origine soient ramenés à l (une réalisation de L). Soit G la variable définie sur le groupe l exactement de la même manière que F l'a été au § 1.2.1.1. pour le groupe k. Il est évident que la distribution conditionnelle de G (la condition étant $L = l$) est une distribution de Fisher-Snedecor avec $(l - 1)$ et et $(n' - 1)$ degrés de liberté (n' étant l'effectif total de l'essai). Par conséquent,

$$P\{G \geqslant g\} = \sum_{l=2}^{k} P_k\{L = l\} . P\{F(l-1, n'-l) \geqslant g\} ,$$

où g est un nombre quelconque, par exemple la réalisation de G sur les échantillons. Généralement, les tables de la distribution de Fisher-Snedecor ne donnent pas $P\{F(l-1, n'-1) \geqslant g\}$ pour les valeurs de $(l-1)$, $(n'-1)$ et g observées dans la pratique. Mais en vertu d'une remarque faite au § 1.2.1.1 on peut également écrire

$$P\{G \geqslant g\} = \sum_{l=2}^{k} P_k\{L = l\} . I_{x_l}\left(\frac{n'-l}{2}, \frac{l-1}{2}\right)$$

où $x_1 = (n'-l)/[n'-l+(l-1)g]$.

Reprenons dans le même ordre les deux exemples du § 1.2.1.1. Supposons tout d'abord que l'on teste H_o contre $H_p : \mu_C > \mu_B > \mu_A$. Dans ce cas, comme les moyennes échantillons sont dans l'ordre de H_p, $l = k = 3$ et la réalisation de G est $g = f = 21$. On calcule facilement, grâce aux tables de la fonction bêta incomplète (Pearson, 1956) et à la table 1.5, la probabilité pour que G soit égal à ou dépasse 21, soit

$$P\{G \geqslant g\} = P_3\{L = 2\} . I_{x_2}\left(8, \frac{1}{2}\right) + P_3\{L = 3\} . I_{x_3}\left(\frac{15}{2}, 1\right) = 162 . 10^{-6} .$$

On prend donc la décision D en faveur de H_p.

Supposons ensuite, toujours pour le même exemple, que l'on veuille tester H_o contre $H_p : \mu_C > \mu_A > \mu_B$. Cette fois, les moyennes observées $\overline{y}_A$ et $\overline{y}_B$ sont dans l'ordre

opposé à celui de H_p. On réunit donc les échantillons correspondants et l'on constitue ainsi un nouvel échantillon (AB) d'effectif 12 et de moyenne 10,5 (inférieure à $\overline{y}_C$). L'analyse de la variance, pour les deux échantillons (AB) et (C) fournit les résultats suivants :

$$c = (216)^2/18 = 2592 \; ; \; s_T = 114 \; ; \; s_t = 81 \; ;$$

$$s_e = 33 \; ; \; v_t^2 = s_t = 81 \; ; \; v_e^2 = s_e/16 = 2{,}0625 \; ;$$

et finalement, $g = v_t^2/v_e^2 = 39{,}27$. La probabilité pour G d'être égal à ou de dépasser g est

$$P\{G \geqslant g\} = P_3\{L = 2\} \cdot P\{F(1,16) \geqslant g\} + P_3\{L = 3\} \cdot P\{F(2,15) \geqslant g\}$$

$$= \frac{1}{2} I_{x_2}\left(8, \frac{1}{2}\right) + \frac{1}{6} I_{x_3}\left(\frac{15}{2}, 1\right)$$

où $x_2 = 0{,}29$ et $x_3 = 0{,}16$. Donc $P\{G \geqslant g\} = 6 \cdot 10^{-6}$. On prend la décision D en faveur de H_p. Si l'on confronte ce résultat avec le précédent, il faut conclure que le traitement C se distingue des deux autres eux-mêmes indistingables.

Considérons le second exemple du § 1.2.1.1 et testons H_o contre H_p : $\mu_C > \mu_B > \mu_A$. Les moyennes observées $\overline{y}_A = 10$, $\overline{y}_B = 11$ et $\overline{y}_C = 12{,}16$ sont dans l'ordre de H_p. La valeur calculée de g est donc celle obtenue pour f, à savoir, 3,5. Et

$$P\{G \geqslant 3{,}5\} = \frac{1}{2} I_{x_2}\left(8, \frac{1}{2}\right) + \frac{1}{6} I_{x_3}\left(\frac{15}{2}, 1\right)$$

où $x_2 = 0{,}8205$ et $x_3 = 0{,}6818$. Par conséquent, $P\{G \geqslant 3{,}5\} = 0{,}0493$. On est donc amené cette fois à prendre la décision D, c'est-à-dire, à rejeter H_o en faveur de H_p. Rappelons que le test-F de H_o contre H_g avait conduit à prendre la décision D_o. On voit que le test-G de H_o contre H_p rejette H_o (en faveur de H_p). Le caractère particulier des alternatives rend le test-G plus « exigeant » en ce qui concerne H_o ; ceci illustre une remarque faite au § 1.1.a.

1.3. *Populations normales hétéroscédastiques. Hypothèse d'homogénéité relative contre alternatives générales.*

Lorsqu'on n'a aucune raison impérative de croire que les populations étudiées sont non-normales, mais que les variances populations sont vraisemblablement inégales et que les effectifs échantillons sont différents, on peut, pour tester H_o contre H_g, utiliser la méthode suivante, due à Welch (1951). Une méthode analogue due à James (1951), a également été proposée ; nous n'en parlerons pas : les deux méthodes n'ont pas été comparées systématiquement.

L'hétérogénéité des variances populations peut être testée de différentes manières. L'une d'elles est le test - χ^2 dû à Bartlett (voir, par exemple, Hald, 1955, p. 290). Box (1953), a montré que, dans l'ignorance où l'on se trouve généralement quant aux fonctions de distribution des populations étudiées, l'application préalable d'un test d'homogénéité des variances peut conduire à des mécomptes plus graves encore que ceux qui peuvent advenir si l'on s'en passe. D'après Box (1953), l'application d'un tel test, préalablement à l'analyse de la variance ou la méthode de Welch, « is rather like putting to sea in a rowing boat to find out whether conditions are sufficiently calm for an ocean liner to leave port ! »

Dans le cas où les fonctions de distribution des populations étudiées peuvent être supposées normales ou approximativement telles, on adopte l'attitude suivante :

— si les effectifs échantillons sont égaux ou presque, on applique le test-F exposé au § 1.2.1.1 ;

— si les effectifs échantillons sont très inégaux et si les variances populations sont différentes, on applique la méthode de Welch (1951), exposée ci-dessous.

Soit π_i la population de toutes les réponses possibles au i^e traitement ($i = 1, 2, \ldots, k$). La fonction de distribution de π_i est $N(\mu_i, \sigma_i)$. Soit Y_{ij} la réponse de la j^e unité expérimentale ayant subi le i^e traitement. On suppose, comme au § 1.2.1.1, que $Y_{ij} = \mu_i + X_{ij}$ où X_{ij} est une variable $N(O, \sigma_i)$; μ_i ($= \mu + \tau_i$) est donc la moyenne de la population π_i. L'hypothèse que l'on teste, H_o : $\mu_1 = \mu_2 = \ldots = \mu_k$ est une hypothèse d'homogénéité relative (voir § 1.1.d). Les alternatives sont générales. Les variances populations étant différentes il ne peut être question d'une estimation de la forme $\sum_{ij}(y_{ij} - \overline{y_{ij}})^2$.

Il faut estimer chaque variance population σ_i^2 au moyen de la variance échantillon

$$S^2_i = \frac{1}{n_i - 1} \sum_j (Y_{ij} - \overline{Y}_{i.})^2 .$$

Posons

$$W_i = n_i / S^2_i ,$$

au moyen de quoi on définit une moyenne pondérée (relativement à l'hétérogénéité des variances)

$$\overline{\overline{Y}} = \frac{\Sigma W_i \overline{Y}_{i.}}{\Sigma W_i}$$

En suite de quoi on peut définir une variation totale (pondérée) due aux traitements

$$\overline{\overline{S}}_t = \sum_i W_i (\overline{Y}_{i.} - \overline{\overline{Y}})^2 ,$$

On définit aussi

$$S_v = \sum_i \frac{1}{n_i - 1} \left(1 - \frac{W_i}{\Sigma W_i}\right)^2$$

Nous verrons plus loin comment il convient de calculer les réalisations de S_t et S_v avec le maximum d'économie. On définit ensuite une variable

$$\overline{\overline{F}} = \frac{\dfrac{1}{k-1} \overline{\overline{S}}_t}{1 + \dfrac{2(k-2)}{k^2 - 1} S_v}$$

dont Welch (1951) a montré qu'elle avait une distribution d'échantillonnage approximativement de Fisher-Snedecor, avec

$$f_1 = k - 1$$

$$f_2 = (k^2 - 1)/3s_v$$

degrés de liberté (s_v est la réalisation de S_v sur l'essai).

Le calcul de $\overline{\overline{s_t}}$ et s_v est relativement simple si l'on utilise les formules suivantes

$$\overline{\overline{s_t}} = \Sigma w_i \overline{y}^2_{i.} + \overline{\overline{y}}^2 \Sigma w_i - 2 \overline{\overline{y}} \Sigma w_i \overline{y}_{i.}$$

$$s_v = \Sigma \frac{1}{n_i - 1} + \frac{1}{(\Sigma w_i)^2} \Sigma \frac{w_i^2}{n_i - 1} - \frac{2}{\Sigma w_i} \Sigma \frac{w_i}{n_i - 1} ,$$

dont chacun des termes s'obtient aisément si l'on aménage les calculs suivant les tableaux 1.6 et 1.7.

Trts i	y_{ij}	n_i	$y_{i.}$	$\bar{y}_{i.}$	$\bar{y}^2_{i.}$	$1/(n_i-1)$	$\sum_j y^2_{ij} - \bar{y}^2_{i.}$	s^2_i	w_i
1									
2									
⋮									
		n'	$y_{..}$			$\sum \frac{1}{n_i-1}$			$\sum w^i$

Table 1.6. Données et calcul de w_i et Σw_i .

Trts i	$w_i \bar{y}_{i.}$	$w_i \bar{y}^2_{i.}$	$w_i/(n_i-1)$	w^2_i	$w^2_i/(n_i-1)$
1					
2					
⋮					
	$\Sigma w_i \bar{y}_{i.}$	$\Sigma w_i \bar{y}^2_{i.}$	$\Sigma \frac{w_i}{n_i-1}$		$\Sigma \frac{w_i^2}{n_i-1}$

Table 1.7. Calcul de $\bar{\bar{s}}_t$ et s_0 .

L'exemple suivant est emprunté à Welch (1951). Nous l'avons quelque peu modifié afin d'illustrer le fait suivant : Les variances échantillons étant très différentes le test-F n'est pas applicable ; si malgré tout, il est appliqué, il conduit à un résultat différent de celui auquel conduit le test-$\bar{\bar{F}}$. Il est d'ailleurs vraisemblable que la probabilité de prendre la décision D_0 avec le test-F est inférieure à la probabilité de prendre la décision D_0 avec le test-$\bar{\bar{F}}$, lorsque les moyennes populations sont égales.

L'exemple est relatif à trois populations π_1 , π_2 et π_3 dans lesquelles on a prélevés des échantillons d'effectifs 20 , 10 , 10 , de moyennes respectives 27,845 , 24,1 , 22,2 et variances (s_i^2) 60,1 , 6,3 , 15,4 . Les données sont rassemblées au tableau 1.8.

Trts i	n_i	$\bar{y}_{i.}$	$\bar{y}^2_{i.}$	s^2_i	w_i	$\frac{1}{n_i-1}$
1	20	27,8454	775,3663	60,1	0,333	0,0526
2	10	24,1	580,810	6,3	1,587	0,1111
3	10	22,2	492,840	15,4	0,649	0,1111
	40				2,569	0,2748

Table 1.8. Données et calcul de w_i .

Trts i	$w_i \bar{y}_{i.}$	$w_i \bar{y}_{i.}^2$	$\frac{w_i}{n_i - 1}$	w_i^2	$\frac{w_i^2}{n_i - 1}$
1	—	—	—	0,1109	—
2	—	—	—	2,5186	—
3	—	—	—	0,4212	—
	61,9267	1499,7956	0,2659		0,3324

Table 1.9. Calcul de $\bar{\bar{s}}_t$ et s_v .

Le test-F est tout d'abord réalisé. On obtient successivement $c = 26.005{,}31$; $s_t = 238{,}52$; $s_e = 1337{,}20$; $v_t^2 = 119{,}262$; $v_e^2 = 36{,}140$; et finalement $f = 3{,}3$. Or, la valeur du seuil 5 % de la distribution de Fisher-Snedecor avec $k - 1 = 2$ et $n' - k = 37$ degrés de liberté est 3,255. Par conséquent, on prend la décision D en faveur de H_g .

Les variances échantillons étant très différentes, les effectifs échantillons étant très différents, les populations étant supposées normales, l'on se trouve dans les conditions d'application du test-$\bar{\bar{F}}$. Les calculs de w_i , $\bar{\bar{s}}_t$ et s_v sont résumés aux tableaux 1.8 et 1.9 et complétés par

$$\bar{\bar{v}} = 24{,}1054 \; ; \; \bar{\bar{y}}^2 = 581{,}0689 \; ; \; (\Sigma w_i)^2 = 6{,}5998 \; ;$$

et finalement $\bar{\bar{s}}_t / 2 = 3{,}513$; d'autre part, $1 + \frac{6}{8} s_v = 1{,}029$. Donc, $\bar{\bar{f}} = 3{,}35$. Or, la valeur du seuil 5 % de la distribution de Fisher-Snedecor avec $k - 1 = 2$ et $(k^2 - 1)/3\, s_v = 22{,}56$ degrés de liberté, est 3,43. Par conséquent, on prend la décision D_o .

1.4. *Populations non-normales.*

1.4.1. *Hypothèse d'homogénéité contre alternatives générales.*

Soient π_i , $(i = 1, 2, \ldots, k)$, k populations dont les fonctions de distribution sont les fonctions $F_i(x)$ dont les dérivées sont continues. On ne sait rien de la forme des fonctions F_i et en particulier, on ne peut pas supposer que ce sont des fonctions de distribution normales.

Soit $F(x)$ une fonction de distribution dont la forme analytique est inconnue. Supposons que la moyenne de la population dont $F(x)$ est la fonction de distribution soit μ et que son écart-type soit σ . Les fonctions de distribution F_i ont toutes la même forme analytique que F , mais diffèrent l'une de l'autre par la valeur d'un paramètre de position. On a

$$F_i(x) = F(x - v_i) \; ,$$

où v_i représente l'effet du i^e traitement. La valeur moyenne de la population dont F_i est la fonction de distribution est $\mu + v_i$. Donc, plus v_i est grand, plus la moyenne du i^e traitement est élevée. D'autre part, on suppose que l'effet d'un traitement ne se manifeste que sur la position de la fonction de distribution. Cela signifie, par exemple, qu'un traitement n'est en aucune manière caractérisé par une valeur de dispersion. Toutes les fonctions de distribution F_i ont le même écart-type σ .

Si $v_1 = v_2 = \ldots = v_k$, les populations π_i sont identiques ; les traitements ne se distinguent pas l'un de l'autre. Par conséquent, l'égalité des v_i implique l'homogénéité des populations et constitue l'hypothèse nulle H_o qu'il faut tester. Nous désignons par D_o la décision statistique de considérer H_o comme l'hypothèse vraie. Les alternatives à H_o

expriment qu'au moins un ν_i se distingue des autres. Il y a donc $(2^k - 1)$ alternatives. Parmi elles, considérons H_k : $\nu_1 = \nu_2 = \ldots \nu_{k-1} = 0 = \nu_k - \Delta$ où Δ est un nombre positif. Nous supposons que

$$\Delta = (\mu + 1)/\sigma \sqrt{m} ,$$

où $m \geqslant 1$. L'hypothèse H_k exprime que les moyennes des traitements sont telles que

$$\mu_1 = \mu_2 = \ldots = \mu_{k-1} = \mu ,$$

$$\mu_k = \mu + \Delta ,$$

m étant un paramètre variable auquel l'expérimentateur peut donner n'importe quelle valeur supérieure à un. Nous désignons par D_k la décision statistique de considérer H_k comme l'hypothèse vraie.

Soit $X_{i1}, X_{i2}, \ldots, X_{in_i}$ l'échantillon aléatoire prélevé dans π_i .

Suivant une notation, reprise au § 3.3 , soit $< X_1 - X_2 - \ldots - X_k >$ la suite ordonnée par valeurs croissantes de toutes les variables observées.

Chaque variable X_{ij} , $(j = 1, 2, \ldots, n_i)$, a dans la suite ordonnée une certaine place, un certain rang R'_{ij}. Soit $R_i = \sum_j R'_{ij}$ la somme des rangs des variables de l'échantillon prélevé dans π_i . Posons également $n' = \sum_i n_i$, l'effectif total de l'expérience.

Considérons la variable aléatoire

n_1	n_2	n_3	h_α	$10^3\, P\{ H > h_\alpha \}$	n_1	n_2	n_3	h	$10^3\, P\{ H > h_\alpha \}$
2	2	2	457	67	5	2	1	500	48
3	2	2	471	48	5	2	2	504	56
3	3	1	514	43	5	3	1	487	52
3	3	2	514	61	5	3	2	525	49
3	3	3	560	50	5	3	3	534	50
4	2	1	482	57	5	4	1	486	56
4	2	2	512	52	5	4	2	527	50
4	3	1	521	50	5	4	3	563	50
4	3	2	540	52	5	4	4	562	50
4	3	3	573	50	5	5	1	491	53
4	4	1	496	48	5	5	2	525	51
4	4	2	524	52	5	5	3	563	51
4	4	3	558	50	5	5	4	564	50
4	4	4	569	49	5	5	5	566	51

Table 1.10. — Valeurs de h_α telles que $P\{ H < h_\alpha \} \cong 0{,}95$.

$$H = \frac{12}{n'(n'+1)} \sum_1^k \frac{R_i^2}{n_i} - 3(n'+1) .$$

La variable H a été proposée par Wallis et Kruskal, (1952), en remplacement de la variable F de Fisher-Snedecor utilisable seulement dans le cas de distributions normales homoscédastiques. Les propriétés du test de H_o basé sur H ont été étudiées par Andrews, (1954). La distribution d'échantillonnage, lors H_o est vrai, a été ajustée récemment par Wallace, (1959).

On prend la décision D_o si

$$H < h_\alpha$$

où h_α est un nombre tel que $P\{D_o|H_o\} = 1 - \alpha$. Lorsque $k = 3$ et $n_i \leqslant 5$, les valeurs de h_α telles que $P\{D_o|H_o\}$ soit voisin de 0,95 sont fournies à la table 1.10, empruntée à Wallis et Kruskal (1952). Lorsque $k = 3$ et $n_i > 5$, ou bien, lorsque $k > 3$, on a

$$P\{D_o \mid H_o\} = P\{\chi^2_{k-1} < h_\alpha\}$$

et dans ces conditions, h_α est le seuil α % de la distribution χ^2_{k-1} (distribution χ^2 avec $(k-1)$ degrés de liberté). Par exemple, si $\alpha = 0{,}05$ et $k = 7$, $h_\alpha = 12{,}6$.

Supposons que l'on ait des raisons de craindre que la population π_k se distingue seule des autres. Dans ce cs, l'alternative à H_o dont il convient de tenir compte est H_k. Le test doit être tel que la probabilité de prendre la décision D_k si H_k est vraie, soit contrôlable et fixée à un niveau élevé. Nous avons considéré ce problème dans le cas où tous les effectifs échantillons sont égaux à n. Il convient de fixer n à l'avance de telle sorte que $P\{D_k|H_k\} \geqslant 0{,}90$. On démontre qu'il faut que

$$n > \frac{\lambda k}{\Delta^2 (k-1)},$$

où λ dépend de k et est fourni à la table 1.11.

k	3	4	5	6	7
λ	13	14	15	16	17

Table 1.11. — Valeurs de λ correspondant à différentes valeurs de k et telles que $P\{D_k|H_k\} \geqslant 0{,}90$, lorsque $\alpha = 0{,}095$.

Dans l'exemple qui suit, le calcul de H n'est pas gêné par la présence d'observations égales dans la suite ordonnée $< X_1 - X_2 - \ldots - X_k >$. Si cette circonstance survient, on remplace le rang de chaque observation double (ou triple, ou quadruple, etc.) par la moyenne de leurs rangs.

	1er trait.		2e trait.		3e trait.		4e trait.		
	obs.	rangs	obs.	rangs	obs.	rangs	obs.	rangs	
	— 2	1	— 1,95	2	— 1,9	3	1	13	
	— 0,8	4	— 0,7	5	— 0,5	6	1,5	14	
	0	7	0,1	8	0,2	9	2,2	18	
	0,3	10	0,4	11	0,5	12	2,5	19	
	2,0	15	2,1	16	2,15	17	3	20	
R_i		37		42		47		84	210
R_i^2		1369		1764		2209		7056	
R_i^2/n		274		353		442		1411	2480

Supposons qu'un expérimentateur doive comparer quatre traitements. Il ne peut supposer que les populations des réponses sont normales ; mais l'expérience lui enseigne que les populations ne diffèrent que par la valeur d'un paramètre central, la moyenne ou la médiane. L'un des traitements, le 4e par exemple, est suspecté d'avoir une valeur contrale supérieure à celle des trois autres ; il convient que $P\{D_4|\Delta = 2\} > 0{,}90$. Pour satisfaire à cette condition il faut prendre (μ étant voisin de zéro et σ voisin de 1)

$$\frac{\lambda k}{4(k-1)} = \frac{14.4}{4.3} \cong 5$$

observations dans chaque population. Si les quatre traitements sont identiques, il convient que $P\{D_0|H_0\} = 0{,}95$. Par conséquent, on prend la décision D_0 si $h < 7{,}81$. Les données sont reproduites au tableau ci-contre.

La somme des rangs (210) est indiquée au tableau comme vérification. Il faut en

$$\Sigma R_i = \frac{n'(n'+1)}{2} \left(= \frac{20 \times 21}{2} \right)$$

On indique aussi $\Sigma R_i^2/n = 2480$ et on calcule ensuite la valeur observée de H, soit

$$h = \frac{12.2480}{20.21} - 3.21 \cong 8$$

On est donc amené à ne pas prendre la décision D_0. Et comme le 4e échantillon présente l'anomalie à laquelle on s'attendait de sa part, on prend la décision D_4 ; on a $P\{D_4|H_4\} > 0{,}90$.

1.4.2. *Hypothèse d'homogénéité contre alternatives particulières.*

Soient π_i, $(i = 1, 2, \ldots, k)$, k populations dont les fonctions de distribution $F_i(x)$ inconnues sont non-normales, mais admettent des dérivées continues. On désire tester l'hypothèse d'homogénéité H_0 : $F_1 = F_2 = \ldots = F_k$ contre les alternatives particulières H_p : $F_1 > F_2 > \ldots > F_k$. Le test que nous exposons est une généralisation, proposée par Jonckheere, (1954), du test basé sur le coefficient des rangs de Kendall.

Il importe de bien comprendre le sens des alternatives H_p. Considérons deux populations seulement π_1 et π_2 dont les fonctions de distribution sont F_1 et F_2. Soit également X_1 la variable aléatoire dont les valeurs possibles constituent la population π_1 et soit X_2 la variable aléatoire dont les valeurs possibles constituent la population π_2. Lorsque $F_1 = F_2$, la probabilité pour que X_1 soit supérieure (ou inférieure) à X_2 vaut exactement $\frac{1}{2}$. Par contre, si $F_1 > F_2$, $P\{X_1 > X_2\} < \frac{1}{2}$ et si $F_1 < F_2$, $P\{X_1 > X_2\} > \frac{1}{2}$. Si $F_1 > F_2$, on a aussi que $\nu_1 < \nu_2$, ν_1 et ν_2 étant respectivement les médianes des populations π_1 et π_2 ; si $F_1 < F_2$, $\nu_1 > \nu_2$.

Généralisant ce qui précède, appelons X_i la variable aléatoire dont les valeurs possibles constituent la population π_i. Et soit N_{ij} la variable aléatoire représentant le nombre de fois que $X_i < X_j$. La variable aléatoire sur laquelle le test de H_0 est basé est

$$S = 2 \sum_{i<j} \sum N_{ij} - \sum_{i<j} \sum n_i n_j ,$$

où n_i *et* n_j sont respectivement les effectifs des échantillons prélevés dans π_i et π_j. Considérons un exemple simple qui permet d'illustrer le calcul d'une réalisation s de S. Nous avons reproduit au tableau ci-contre les observations prélevées dans trois populations π_1, π_2 et π_3 ; nous les avons ordonnées par valeurs croissantes.

π_1	π_2	π_3
1	3	5
4	6	8
9	11	13

Dans cet exemple, $n_1 = n_2 = n_3 = 3$. La deuxième somme intervenant au second membre de S vaut donc

$$\sum_{i<j} \sum n_i n_j = n_1 n_2 + n_1 n_3 + n_2 n_3 = 27 .$$

De même, la première somme vaut

$$2 \sum_{i<j} \sum n_{ij} = 2 (n_{12} + n_{13} + n_{23}) ,$$

où n_{12} est le nombre de fois qu'une observation de π_1 est inférieure à une observation de π_2, n_{13} et n_{23} étant définis de la même manière. Par conséquent, $n_{12} = 3 + 2 + 1 = 6$, $n_{13} = 3 + 3 + 1 = 7$ et $n_{23} = 3 + 2 + 1 = 6$. Donc $2 \sum_{i<j} \sum n_{ij} = 38$ et $s = 11$.

Lorsque H_o est vrai, la variable X_i n'a pas plus de chance de dépasser la variable X_j que de lui être inférieure. Par conséquent, la valeur moyenne de N_{ij}, lorsque H_o est vrai, est $n_i n_j/2$ et la valeur moyenne de S vaut zéro. Par contre, lorsque H_p est vrai, N_{ij} a tendance à prendre des valeurs supérieures à $n_i n_j/2$ et S a donc tendance à prendre des valeurs positives. La valeur maximum de S est $\sum_{i<j} \sum n_i n_j$. Les grandes valeurs de S sont donc significatives d'un écart par rapport à H_o en faveur de H_p.

Posons

$$\sigma^2 = \frac{1}{18} [n'^2 (2n' + 3) - \Sigma n_i^2 (2n_i + 3)] ,$$

où $n' = \Sigma n_i$; si tous les effectifs échantillons sont égaux à n, on a $n' = kn$ et

$$\sigma^2 = \frac{1}{18} k (k-1) n^2 [2n (k+1) + 3] .$$

Soit D_o la décision statistique de considérer H_o comme vraie et soit D_p la décision statistique de considérer H_p comme vraie. La règle que nous donnons ci-dessous est telle que $P\{ D_o \mid H_o \} \geqslant 0{,}95$. Les nombres du tableau 1.12 permettent de l'appliquer dans le cas où tous les effectifs échantillons sont égaux ($\leqslant 10$) et où $k \leqslant 6$. On prend la décision D_o si s est inférieur au nombre de la table 1.12 correspondant à n et à k ; on prend la décision D_p si s lui est supérieur ou égal. Pour les configurations d'effectifs qui ne sont pas prévues à la table 1.12, on prend la décision D_o si $s < 1{,}64\,\sigma$ et la décision D_p si $s \geqslant 1{,}64\,\sigma$, où σ est la racine carrée positive de σ^2.

n \ k	3	4	5	6
2	10°	14°	20°	26°
3	17°	26°	34°	43
4	24°	38°	49	65
5	33°	49	69	90
6	41	64	90	118
7	51	80	113	149
8	62	97	138	173
9	74	116	164	216
10	86	136	192	253

Table 1.12. Seuils de signification 5 % pour le test - S . Les nombres marqués de (°) sont empruntés à Jonckheere (1954) ; les autres sont obtenus grâce à l'approximation normale indiquée par cet auteur qui fait en outre remarquer que l'approximation est pratiquement très bonne même dans les cas extrêmes où les effectifs sont très différents.

Appliquons la méthode à un exemple emprunté à Jonckheere (1954) ; les données sont rassemblées au tableau 1.13. Les quatre populations

	π_1	π_2	π_3	π_4
	19	21	40	49
	20	61	99	110
	60	80	100	151
	130	129	149	160
$x_{i.}$	229	291	388	470
$\overline{x_{i.}}$	57,25	72,75	97	117,50
s_i^2	2037,69	1508,19	1491,5	1919,25

Table 1.13. Données de l'exemple.

π_1, π_2, π_3 et π_4 ont des fonctions de distribution non-normales. Les échantillons ont tous le même effectif $n = 4$. Les variances échantillons sont fort différentes. Nous sommes donc dans les conditions d'application des méthodes indépendantes de l'hypothèse de normalité : soit le test - H (§ 1.4.1), soit le test - S. Le choix entre les deux tests s'établit sur la base des alternatives à l'hypothèse d'homogénéité : si l'on teste H_o contre H_g, on utilise le test - H ; si l'on teste H_o contre H_p, on utilise le test - S. A titre d'illustration, nous appliquons aux données de la table 1.13 successivement le test - F (seulement valable si l'on suppose la normalité des populations), le test - H et le test - S.

Le test - F (voir § 1.2.1.1.) conduit à la valeur observée $f = 1{,}216$; or, le seuil 5 % de la distribution de Fisher-Snedecor avec 3 et 12 degrés de liberté est 3,49. On prend donc la décision D_o (contre la décision D_g).

Le test - H (voir § 1.4.1) conduit à la valeur observée $h = 3{,}73$; or, le seuil 5 % de la distribution de Fisher-Snedecor avec 3 et 12 degrés de liberté est 3,49. On prend donc la décision D_o (contre la décision D_g).

Les réalisations de la variable aléatoire N_{ij} sont $n_{12} = 11$, $n_{13} = 12$, $n_{14} = 13$, $n_{23} = 11$, $n_{24} = 12$, et $n_{34} = 12$; donc $s = 46$. Or, le seuil 5 % correspondant à $n = 4$ et $k = 4$ (voir table 1.12) est 38. On prend donc la décision D_p (contre la décision D_o). On rejette l'hypothèse d'homogénéité en faveur de l'alternative H_p : $F_1 > F_2 > F_3 > F_4$. Autrement dit, on rejette l'hypothèse d'égalité des médianes contre l'aternative $\nu_1 < \nu_2 < \nu_3 < \nu_4$; mais il est bien entendu que le test - S est relatif à l'hypothèse H_o d'homogénéité totale ; la formulation des conclusions en termes de médianes n'est qu'un aspect de la question : si l'on avait pris la décision D_o, cela aurait signifié que l'on acceptait l'hypothèse que les quatre populations sont en tout point identiques (et pas seulement de médianes égales).

Comme l'exemple 1.2.2., celui-ci illustre la même remarque faite au § 1.1.a sur la plus grande efficacité d'un test de H_o contre H_p.

REFERENCES BIBLIOGRAPHIQUES POUR LE CHAPITRE 1

ANDREWS, C. (1954). — *Asymptotic behavior of some rank tests for analysis of variance*. A.M.S., pp. 724-735.

BARTHOLOMEW, D.J. (1959). — *A test of homogeneity of ordered alternatives*. Bka, I, pp. 36-48, II, pp. 328-335.

BOX, G.E.P., (1953). — *Non-normality and tests on variances*. Bka, pp. 318-335.

DAVID, F.N. & JOHNSON, N.L. (1951). — *The effect of non-normality on the Power function of the F-test in the analysis of variance*. Bka, pp. 43-57.

DE MUNTER, P. (1954). — *Sur les transformations de variables aléatoires*. « Bull. Soc. Belge de St. », pp. 99-111.

DE MUNTER, P. (1960). — *Sur la notion de traitement et le problème de la comparaison de plus de deux traitements*. « Bull. Inst. Agr. Gembloux » (centenaire).

FELDT, L.S. & MAHMOUD, M.W. (1958). — *Power function charts for specifying numbers of observations in analysis of variance of fixed effects*. A.M.S., pp. 871-877.

HACK, H.R.S. (1958). — *An empirical investigation into the distribution of the F-ratio in samples from two non-normal populations*. Bka, pp. 260-265.

HALD, A. (1955). — *Statistical Theory with Engineering Applications*. J. Wiley, N.Y.

HORSNELL, G. (1953). — *The effect of unequal group variances on the F. test for the homogeneity of group means*. Bka, pp. 128-136.

JAMES, G.S. (1951). — *The comparison of several groups of observations when the ratios of the population variances are unknown*. Bka, pp. 324-329.

JOHNSON, N.L. (1953). — *Some notes on the application of sequential methods in the analysis of variance*. A.M.S., pp. 614-623.

JONCKHEERE, A.R. (1954). — *A distribution-free k-sample test against ordered alternatives*. Bka, pp. 133-145.

KRUSKAL, W.H. & WALLIS, W.A. — *Use of ranks in one-interion variance analysis*. 1952 : J.A.S.A., pp. 583-621 ; 1953 : J.A.S.A., pp. 907-911.

MILES, R.E. (1959). *The complete amalgamation into blocks, by weighted means, of a finite set of real numbers*. Bka, pp. 317-327.

PEARSON, E.S. & HARTLEY, H.O. (1951). — *Charts of the power function of analysis of variance tests, derived from the non-central F-distribution*. Bka, pp. 112-130.

PEARSON, K. (1956). — *Tables of the Incomplete Beto-Function*. University Press, Cambridge.

RAY, W.D. (1956). — *Sequential analysis applied to certain experimental designs in the analysis of variance*. Bka, pp. 388-399.

WALLACE, D.L. (1959). — *Simplified beta-approximations to the Kruskal-Wallis H-test*, J.A.S.A., pp. 225-230.

WELCH, B.L. (1951). — *On the comparison of several mean values : an alternative approach*. Bka, pp. 330-336.

Cycle de cours sur les "Méthodes de contrôle de réception par échantillonnage„

En collaboration avec le Centre de Productivité d'Anvers (A.C.P.) et sous les auspices de la Société Royale des Ingénieurs Flamands (K.V.I.V.), la Fédération Economique Flamande (V.E.V.) et la Fédération Flamande de Chimie, l'A.B.S.I. (Association belge pour les Applications Industrielles de la Statistique) a organisé l'hiver dernier, en langue néerlandaise, un cycle traitant du « Contrôle de Réception ».

Ce cycle suivait une série d'exposés présentés l'an dernier par les première et dernière associations citées ci-dessus, et donnait un aperçu des diverses possibilités d'application de la statistique dans l'industrie.

Comme ni la série complète des conférences ni le nom des orateurs n'ont été précisés dans notre revue de septembre 1960, nous en donnons ci-dessous la nomenclature :

1. — *Le 16 novembre : Soirée préliminaire d'information.*
   - — Introduction par M. DE GRANDE, Président de l'A.B.S.I., Chef de Service à la Firme GEVAERT.
   - — « Que peut attendre le Chef d'entreprise du contrôle statistique de qualité en général et du contrôle scientifique de réception en particulier ? » par M. FIERENS, ir., Administrateur-Directeur de la Firme GEVAERT.
   - — « En quoi consiste le Contrôle de Réception ? Quelles en sont les Possibilités pratiques ? » par M. DE WOLF, statisticien à la Firme GEVAERT.

2. — *Le 20 novembre : Première leçon.*
   « Les distributions en probabilité et leurs propriétés », par M. A. HEYVAERT, Directeur à l'Institut d'Organisation Industrielle et Commerciale (I.O.I.C.).

3. — *Le 14 décembre : Deuxième leçon.*
   « Théorie de l'échantillonnage », par E. DE WOLF, Dr. Sc.

4. — *Le 4 janvier : Troisième leçon.*
   « Contrôle de Réception par Attributs - Echantillonnage simple », par E. DE WOLF.

5. — *Le 18 janvier : Quatrième leçon.*
   « Echantillonnage multiple et séquentiel - Contrôle par attributs », par E. BEFAHY, L. Sc., Secrétaire A.B.S.I., Métrologiste Principale au Ministère des Affaires Economiques.

6. — *Le 1er février : Cinquième leçon.*
   « Contrôle par mesures », par M. MARTENS, ir., Statisticien à la SORCA.

7. — *Le 15 février : Sixième leçon.*
   « Contrôle de qualité par des caractéristiques purement quantitatives », par E. de WILDE, Statisticien à COCKERILL-OUGREE.

8. — *Le 1er mars : Septième leçon.*
   « Contrôle des productions continues », par E. BEFAHY.

9. — *Le 15 mars. Huitième leçon.*
   « Le contrôle de réception et le choix des plans dans la pratique par le Prof. Dr. H.C. HAMAKER, Statisticien en Chef de la S.A. PHILIPS à Eindhoven.

10. — *Le 29 mars. Neuvième leçon.*
    « Problèmes psychologiques posés par l'introduction de nouvelles techniques » par le Prof. Ch. MERTENS de WILMARS, Professeur à l'Université de Louvain.

La soirée d'information fut suivie par une quarantaine de personnes et jugée suffisamment intéressante pour que dix-sept inscriptions émanant pour la plupart d'industries diverses puissent être enregistrées.

L'intérêt fut soutenu de la première à la dernière séance, en effet on a noté seize présences à chaque réunion.

Il est encourageant de constater que tant d'entreprises de la région anversoise et même une firme de Gand ont compris la nécessité de donner à au moins un membre de leur cadre l'occasion de perfectionner leurs connaissances dans le domaine du contrôle par échantillonnage.

Ceci prouve que, finalement, nombreux sont ceux qui, comme nous, ont acquis la conviction que le nouvel instrument qu'est la statistique dans l'industrie, doit être considéré comme indispensable si nous voulons nous introduire ou nous maintenir, grâce à un niveau de qualité valable, dans un Marché commun en plein développement.

Les réactions, que nous avons pu recueillir durant le cycle, montrent clairement que pour plus d'un, les techniques du contrôle de réception ne constituaient encore qu'un terrain en friche. De ce fait, comme d'ailleurs lors du cycle précédent, la partie théorique fut pour d'aucuns assez difficilement assimilable. Heureusement l'avant-dernière séance, présentée par le Professeur HAMAKER préconisa la simplicité et l'uniformité des méthodes employées. Les participants purent en retirer une leçon très pratique, à savoir que, en réalité, une connaissance technique approfondie des méthodes de réception est souhaitable si l'on veut effectuer un choix judicieux et peut-être surtout si l'on désire que la simplification recherchée soit pleinement justifiée théoriquement.

Au cours de la dernière leçon, les auditeurs s'entendirent exposer les problèmes humains et psychologiques que pose l'introduction de nouvelles techniques. Le Professeur Ch. MERTENS de WILMARS donna un aperçu des divers procédés permettant de contrer la trop connue résistance qui s'oppose à tout changement.

Il convient enfin de féliciter pour leur persévérance, les inscrits qui furent infatigablement au poste. Il reste à leur souhaiter un plein succès dans l'introduction et l'emploi des nouvelles techniques et des connaissances qu'ils ont acquises.

Lessencyclus over afnamecontrole

In samenwerking met het Antwerps Centrum voor Productiviteitsbevordering (A.C.P.) en onder de auspiciën van de Koninklijke Vlaamse Ingenieursvereniging (K.V.I.V.), het Vlaams Economisch Verbond (V.E.V.) - Afdeling Antwerpen, en de Vlaams Chemische Vereniging, heeft de Belgische Vereniging voor Industriële Toepassingen van de Statistiek (A.B.S.I.) tijdens het voorbije winterseizoen een cyclus in het vlaams ingericht over « Afnamecontrole ».

Deze cyclus volgde op de reeks voordrachten die vorig jaar doorgingen, eveneens in samenwerking tussen eerst- en laatstsgenoemde verenigingen, en die een overzicht bracht van de verscheidene toepassingsmogelijkheden van de statistiek in de industrie.

Daar noch de volledige reeks der spreekbeurten, noch de sprekers in ons tijdschrift van september 1960 werden opgegeven, laten wij hieronder volledigheidshalve de lijst der sprekers en het door hun behandelde onderwerp volgen :

1. — *16 november : Voorafgaandelijke informatie-avond.*
   - — Inleiding door de heer dr. DE GRANDE, Voorzitter A.B.S.I., Afdelingshoofd bij de N.V. GEVAERT Photo Produkten.
   - — « Wat mag de bedrijfsleiding verwachten van statistische kwaliteitscontrole in het algemeen en van wetenschappelijk verantwoorde afnamecontrole in 't bijzonder », door de heer ir. FIERENS, Beheerder-Direkteur der N.V. GEVAERT Photo Produkten.
   - — « Wat houdt de afnamecontrole in ? Welk zijn de praktische toepassingsmogelijkheden ? » door de heer dr. DE WOLF, Statisticus bij de N.V. GEVAERT Photo Produkten.

2. — *30 november : Eerste les.*
   « Waarschijnlijkheidsverdelingen en hun eigenschappen » door de heer ir. A. HEYVAERT, Direkteur bij het Instituut voor Industriële en Commerciële Organisatie (I.O.I.C.).

3. — *14 december : Tweede les*
   « Steekproeftheorie », door de heer E. DE WOLF.

4. — *4 januari : Derde les.*
   « Afnamecontrole op attributen - Enkelvoudige steekproef », door de heer dr. DE WOLF.

5. — *18 januari : Vierde les.*
   « Meervoudige en sequentiële steekproefsystemen op attributen » door de heer Lic. BEFAHY, Secretaris van A.B.S.I., Eerstaanwezend Metrologist bij het Ministerie van Economische Zaken.

6. — *1 februari : Vijfde les.*
   « Afnamecontrole door meting », door de heer ir. MARTENS, statisticus bij de N.V. SORCA (Toegepaste Economische en Wiskundige Studies).

7. — *15 februari : Zesde les.*
   « Afnamecontrole bij zuiver kwantitatieve kenmerken », door de heer de WILDE, Statisticus bij de N.V. COCKERILL-OUGREE.

8. — *1 maart : Zevende les.*
   « Controle van continue produkties » door de heer Lic. BEFAHY.

9. — *15 maart : Achtste les.*
   « Het gebruik van afnamecontrole en de keuze van steekproefplans in de Praktijk », door de heer Prof. dr. H.C. HAMAKER, Hoofdstatisticus bij de N.V. PHILIPS Gloeilampenfabrieken te Eindhoven.

10. — *29 maart : Negende les.*
    « Psychologische problemen bij het invoeren van nieuwe technieken », door de heer Prof. dr. Ch. MERTENS de WILMARS, Professor aan de Universiteit te Leuven.

De informatieavond werd bijgewoond door circa veertig deelnemers en werd blijkbaar voldoende interessant gevonden opdat op zeventien inschrijvingen, meestal van verschillende industriën afkomstig, zouden toekomen.

De belangstelling bleef gevestigd tot en met de laatste zitting, daar wij gemiddeld zestien aanwezigen mochten noteren.

Het verheugt ons dat zoveel verschillende industriën uit het Antwerpse en zelfs een firma uit Gent de noodzaak aangevoeld hebben om ten minste één lid van haar kader de gelegenheid te geven zich op het gebied van de afnamecontrole de instrueren.

Dit bewijst dat velen uiteindelijk, zoals wij, tot de vaststelling kwamen dat het nieuwe instrument, dat statistiek in de industrie is, als onontbeerlijk moet beschouwd worden willen wij ons in de zich ontwikkelende Euromarkt kunnen vestigen of handhaven dank zij een verantwoorde kwaliteitsconcurrentie.

De reacties die wij tijdens de lessencyclus konden opvangen wezen er op dat voor meerderen de afnamecontrole technieken nog braak terrein betekenden. Van daar dan ook dat, zoals zulks vorig jaar tijdens de vorige cyclus ook het geval was, voor sommigen het theoretisch gedeelte een zwaar te slikken stof. Gelukkig was er in de voorlaatste instantie Prof. dr. H.C. HAMAKER die voor eenvoud en uniformiteit in de gebruikte methodes predikte. Hieruit konden de deelnemers de zeer praktische les trekken dat in de praktijk een diepgaande theoretische kennis van de afnamecontrole technieken wenselijk is opdat een verantwoorde keuze zou kunnen gemaakt worden en ook en misschien vooral op dat als naar vereenvoudiging gestreefd wordt zulks volledig theoretisch zou gemotiveerd zijn.

De deelnemers kregen tevens, dank zij de laatste les, ook een inzicht in de menselijke en psychologische problemen die het invoeren van nieuwe technieken met zich brengen. Prof. dr. C. MERTENS de WILMARS gaf tevens een overzicht van verschillende mogelijke middelen om de klassiek gekende weerstand tegen veranderingen te ontzenuwen.

Ten slotte past het de ingeschrevenen die zo flink steeds op post waren te feliciteren voor hun volharding en hun veel succes toe te wensen bij het invoeren en in gebruik nemen van de nieuwe technieken en kennissen.

Nous attirons tout particulièrement l'attention de nos lecteurs sur le
HUITIEME CONGRES INTERNATIONAL DU T.I.M.S.
(The Institute of Management Sciences)

Ce Congrès se tiendra au Palais des Congrès à Bruxelles, du 23 au 26 août 1961, sous la présidence du Professeur Henri Theil, directeur de l'Institut d'Econométrie de Rotterdam. Le Professeur William Cooper (Graduate School of Industrial Administration, Carnegie Institute of Technology, Pittsburgh 13, Penna, U.S.A.) préside le Comité du Programme et reçoit les textes ou résumés des communications soumises pour présentation au Congrès. Le Comité d'organisation est présidé par le Dr Jacques Drèze, chargé de cours à l'Université de Louvain, 3, avenue Princesse Lydia, Heverlé-Louvain. Les inscriptions sont reçues par M. Max Duval, Office Belge pour l'Accroissement de la Productivité, 60, rue de la Concorde, Bruxelles 5 (600 fr. belges pour les membres du T.I.M.S. — 750 fr. belges pour les non-membres). Le programme du Congrès et les noms des présidents de séances se présentent comme suit :

Mercredi 23 août :

9 h. 30. — Communications diverses I — J. Teghem.
— Communications diverses II — D. Gilford.
12 h. 45. — Déjeuner — allocution à préciser.
14 h. 30. — La Programmation Stochastique — J. Vajda.
— La Simulation et les Jeux — M. Shubik.
17 h. 30. — Réception.

Jeudi 24 août :

9 h. 30. — Décision rationnelle d'investir — G. Kreweras.
— Théorie et Analyse des organisations — G.H. Symonds et J.R. Simpson.
12 h. 45. — Déjeuner — Allocution Présidentielle de M.A. Geisler.
14 h. 30. — Les probabilités subjectives et objectives — H. Solomon.
— Réalisme et théorie dans les sciences de la gestion — C.W. Churchman et R. Crane.
17 h. 30. — Réunion publique du Conseil du T.I.M.S.

Vendredi 25 août :

9 h. 30. — « Adaptive Systems » — R.M. Thrall.
— Sciences du Comportement — R. Bush (à confirmer).
12 h. 45. — Déjeuner — allocution du Professeur Tinbergen : « Country Development Programming ».
14 h. 30. — Rapports des Recherches du Centre International du T.I.M.S. — M. Theil.
— Applications de la Programmation linéaire en nombres entiers — R Gomory.
17 h. 30. — Réunions de Collèges.

Samedi 26 août :

9 h. 30. — Réunions de Collèges.

Les langues officielles du Congrès seront l'anglais et le français. Tous détails complémentaires seront communiqués sur demande adressée à M. Duval.

Le Bureau Universitaire de Recherche Opérationnelle organise en 1961 deux stages d'initiation aux techniques de la Recherche Opérationnelle. Comme les précédents, ces stages sont destinés à aider les entreprises qui sésireraient former quelques spécialistes aux principes et aux méthodes de la Recherche Opérationnelle, mais n'ont pas la possibilité de leur faire suivre les cours ou séminaires échelonnés sur toute une année scolaire. Ils ne visent pas à donner un enseignement complet et suffisant, mais à exposer quelques bases théoriques nécessaires à l'étude des méthodes de Recherche Opérationnelle, et à donner un fil directeur qui permette aux participants de prolonger ces études lorsqu'ils seront rentrés dans leurs entreprises.

Les candidats devront avoir une bonne formation mathématique générale comportant de préférence des notions sérieuses sur le Calcul des Probabilités. Ils n'auront cependant pas besoin de connaissances préalables relatives à la Recherche Opérationnelle.

Chaque stage est réparti sur deux semaines séparées par un mois environ. La première, par le rappel de quelques éléments de mathématiques remet en mémoire des théories et un langage nécessaire à la bonne compréhension des techniques classiques de la Recherche Opérationnelle exposées en ce qu'elles ont d'essentiel dans la deuxième semaine. Le programme ci-joint, bienque sujet à quelques modifications, donne le détail de cet enseignement.

Ces stages auront lieu à Paris selon le calendridr suivant :

Premier stage : 1re semaine du 17 au 21 avril 1961 inclus ; 2e semaine du 29 mai au 2 juin 1961 inclus.

Deuxième stage : 1re semaine du 6 au 10 novembre 1961 inclus ; 2e semaine du 4 au 8 décembre 1961 inclus.

Chaque journée comprendra quatre séances de 1 h. 30 environ. Le nombre des participants sera limité à 20.

Une participation aux frais de 1.000 NF. par personne pour l'ensemble du stage sera demandée aux entreprises.

Les inscriptions sont recueillies au Secrétariat du Bureau Universitaire de Recherche Opérationnelle, 17 rue Richer, Paris 9e.

PROGRAMME DU STAGE D'IMITATION AUX TECHNIQUES DE LA RECHERCHE OPERATIONNELLE

PREMIERE SEMAINE

1. — Grammaire des ensembles fins — Algèbre de Boole — Treillis.
2. — Relations — Correspondances — Fonctions.
3. — Espaces vectoriels — Calcul linéaire.
4. — Quelques problèmes combinatoires et leur représentation algorithmique.
5. — Matrices et formes linéaires — Résolution des équations.
6. — Analyse spectrale (puissances successives d'une matrice carrée).
7. — Opérateurs linéaires : intégrale de Stieltjès.
8. — Calcul des probabilités — Notion de variable aléatoire.
   Liaison en probabilité — Addition de variables aléatoires.
   Epreuves répétées.
9. — Processus linéaires.
10. — Equations fonctionnelles linéaires.
11. — La loi des grands nombres.
12. — L'entropie.
13. — Problèmes de maximums et minimums — Dualité — Inéquations.
14. — Ergodicité d'un processus linéaire.
15. — Espaces fonctionnels : fonctions génératrices et fonctions caractéristiques.
16. — Principales lois de probabilités.
17. — Multiplicateurs de Lagrange — Théorèmes de Kuhn et Tucker.
18. — Processus aléatoires poissonniens.
19. — Processus aléatoires de Markoff.
20. — Discussion générale.

DEUXIEME SEMAINE

21. — Le calcul économique.
22. — La mathématique des programmes économiques.
23. — Production du hasard (nombres au hasard, méthodes de Monte-Carlo).
24. — Les polyèdres et les programmes linéaires.
25. — Les problèmes de la décision en face d'incertitude — Nation d'utilité et probabilités.
26. — La dualité et les programmes linéaires.
27. — Les décisions des statisticiens. Théories classiques des tests, théorie de Wald.
28. — Résolution des programmes linéaires.
29. — Un programme dynamique : résolution par récurrence et résolution par méthode simpliciale.
30. — Files d'attente : La gestion économique du temps perdu.
31. — Quelques problèmes de gestion des stocks.
32. — Exemple d'application à un cas particulier.
33. — Exemple d'application à un cas particulier.
34. — Théorie des jeux.
    1) Théorie mathématique du duel.
35. — Théorie des jeux.
    2) La coopération et la lutte.
36. — Exeple d'application à un cas particulier.
37. — Quelques points fondamentaux de la théorie des programmes dynamiques.
38. — Un exemple de rationalisation des décisions séquentielles fixation des prix sur un marché particulier.
39. — Théorie des jeux.
    3) Les décisions collectives et leurs difficultés logiques (effet Condorcet, théorème d'Arrow).
40. — Discussion générale.

La plupart des conférences sont faites par les membres du Bureau Universitaire de Recherche Opérationnelle : MM. BOUZITAT, GIRAULT, GUILBAUD, KREWERAS, MORLAT, VALETTE.

Nous apprenons avec plaisir que le prix Heuschling vient d'être décerné à M. F. BULTOT, membre de la Société belge de Statistique, chef du Bureau de Climatologie de l'I.N.E.A.C. et membre de l'Académie royale des Sciences d'Outre-Mer.

Ce prix récompense le meilleur ouvrage d'un auteur belge apportant à la statistique une contribution importante. Il est attribué tous les cinq ans par un Jury de neuf membres choisis par le Roi parmi des personnalités présentées par le Conseil supérieur de Statistique, l'Académie royale de Belgique et la « Koninklijke Vlaamse Akademie voor Wetenschappen, Letteren en Schone Kunsten van België ».

Parmi les dernières publications parues aux Editions Eyrolles (61, bd St-Germain, Paris (V^e)), nous avons relevé l'ouvrage suivant :

LE CONTROLE STATISTIQUE DES FABRICATIONS

par René CAVE,

Ancien élève de l'Ecole polytechnique - Ingénieur militaire en chef de l'Armement.

Préface de G. MARMOIS, *Membre de l'Institut.*

Deuxième édition entièrement refondue.

Ce livre a un but pratique. Il donne des méthodes simples d'utilisation en atelier du contrôle statistique. Ainsi, sans entrer dans le détail de la théorie, le personnel de

contrôle pourra comprendre plus facilement l'intérêt et l'utilisation de ces nouvelles techniques.

La refonte complète de l'ouvrage a permis de le compléter par un nombre important de nouveautés et par l'utilisation de notations normalisées.

La première partie fournit au lecteur le matériel statistique nécessaire, c'est-à-dire les définitions des lois principales, mais l'auteur a ajouté diverses considérations pratiques utiles telles que l'interprétation des diagrammes de fréquence, l'utilisation des nombres au hasard, les pertes d'information, etc.

La seconde partie donne les méthodes de contrôle statistique en cours de fabrication par mesures et par calibres (méthodes classiques, méthodes mises au point par l'auteur, méthodes nouvelles et simplifiées récentes). Le travail du contrôleur est facilité par des tableaux en abaques.

La troisième partie contient les indications nécessaires permettant d'établir un contrôle de réception efficace, à prix minimum ou à coût fixé, pour des risques admis par le client et le fournisseur, ou pour une qualité moyenne limite.

La quatrième partie concerne l'application des méthodes statistiques aux recherches ; son contenu est l'outil indispensable permettant de prendre une décision avec des risques d'erreurs réduits au minimum ou tout au moins fixés (comparaison d'appareils, de procédés de fabrication, etc.). Le dernier chapitre précise les activités du Service Qualité.

18 additifs (démonstrations), tableaux récapitulatifs (18 pages), 20 tables (dont 4 nouvelles) et 9 abaques terminent l'ouvrage.

Ainsi rédigé et mis à jour, l'ouvrage de M. Cavé s'adresse non seulement aux statisticiens industriels, mais aussi aux ingénieurs et techniciens intéressés par l'utilisation de ces nouvelles méthodes de contrôle des fabrications.

Wij vragen de speciale aandacht van onze lezers voor het :

ACHTSTE INTERNATIONAAL CONGRES VAN T.I.M.S.

(The Institue of Management Sciences)

Dit congres zal plaats hebben in het Congrespaleis te Brussel, van 23 tot 26 augustus 1961, onder voorzitterschap van Professor Henri Theil, directeur van het Instituut voor Econometrie te Rotterdam, Professor William Cooper (Graduate School of Industrial Administration, Carnegie Institute of Technology, Pittsburgh 13, Penna, U.S.A.) is voorzitter van het programmacomité en ontvangt de teksten of samenvattingen der mededelingen die ingediend worden om op het Congres voorgedragen te worden. Het Organisatiecomité wordt voorgezeten door dr. Jacques Drèze, docent aan de Leuvense Universiteit, Prinses Lydialaan 3, Heverlee-Leuven. De inschrijvingen worden ingewacht door M. Max Duval, Belgische Dienst voor Opvoering van de Produktiviteit, Eendrachtstraat 60, te Brussel (600 B.Fr. voor de leden van T.I.M.S. — 750 B.Fr. voor niet leden). Het programma van het Congres en de namen van de voorzitters der vergaderingen zijn als volgt :

Woensdag 23 augustus :

9 u. 30. — Diverse mededelingen I — J. Teghem.
— Diverse mededelingen II — D. Gilford.
12 u. 45. — Middagmaal — nog nader te bepalen toespraak.
14 u. 30. — Stochastische programmatie — J. Vajda.
— Simulatie en Spelen — M. Shubik.
17 u. 30. — Receptie.

Donderdag 24 augustus :

9 u. 30. — Rationele investeringsbeslissingen — G. Kreweras.
— Theorie en analyse der organisatie — G. H. Symonds et J.R. Simpson.
12 u. 45. — Middagmaal — Toespraak door de voorzitter M. A. Geisler.
14 u. 30. — Subjectieve en objectieve waarschijnlijkheden — H. Solomon.
— Realisme en theorie in de wetenschap van het bedrijfsbeheer — C. W. Churchman en R. Crane.
17 u. 30. — Openbare Beheerraadsvergadering van T.I.M.S.

Vrijdag 25 augustus :

9 u. 30. — « Adaptive Systems » — R. M. Thrall.
— Studie van de gedragingen (onder voorbehoud) — R. Bush.
12 u. 45. — Middagmaal — Toespraak door Prof, Tinbergen « Country Development Programming ».
14 u. 30. — Verslag der opzoekingen van het Internationaal Centrum van T.I.M.S. — M. Theil.
— Toepassingen van Lineaire Programmatie met gehele getallen — R. Gomory.
17 u. 30. — Vergaderingen der werkgroepen.

Zaterdag 26 augustus :

9 u. 30. — Vergaderingen der werkgroepen.

De officiële voertalen van het Congres zijn engels en frans. Alle bijkomende inlichtingen kunnen bekomen worden bij M. Duval.

Het « Bureau Universitaire de Recherche Opérationnelle » van de Université de *Paris* organiseert in 1961 twee stages gewijd aan een inleiding tot de technieken van het Operationeel Onderzoek. Zoals de vroegere, zijn ook deze stages bedoeld als hulp voor de bedrijven die enkele specialisten in de principes en de methodes van het Operationeel Onderzoek wensen te vormen, maar die niet in de mogelijkheid zijn hen cursussen of seminaries te laten volgen die over een heel schooljaar verspreid zijn. Het is niet de bedoeling om een volledig onderwijs te geven, maar veeleer om enkele theoretische grondprincipes uiteen te zetten die onmisbaar zijn bij de studie der methodes van het Operationeel Onderzoek, en om een leidraad te verschaffen die aan de deelnemers moet toelaten hun studie verder te zetten in hun bedrijf.

De kandidaten dienen een stevige wiskundige vorming te hebben die bij voorkeur ernstige noties van waarschijnlijkheidsleer omvat. Er is echter geen enkele voorafgaande kennis van Operationeel Onderzoek vereist.

Elke stage is gespreid over twee weken, telkens gescheiden door ongeveer een maand. De eerste week brengt, door het hernemen van enkele wiskundige beginselen, de theoriën en de taal in herinnering die nodig zijn tot het begrijpen der klassieke technieken van Operationeel Onderzoek. Het wezenlijke van deze laatste wordt dan tijdens de tweede week uiteengezet. Onderstaand programma omvat, behoudens enkele mogelijke wijzigingen, het détail van dit onderricht.

Deze stages zullen plaats hebben te Parijs op de volgende data :

Eerste stage : 1e week : van 17 tot en met 21 april 1961 ; 2e week : van 29 mei tot en met 2 juni 1961.

Tweede stage : 1e week : van 6 tot en met 10 november 1961 ; 2e week : van 4 tot en met 8 december 1961.

Er zullen dagelijks 4 lessen gegeven worden van ongeveer 1 ½ u. Het aantal deelnemers wordt beperkt tot 20.

Per deelnemer wordt vanwege het bedrijf een deelname in de kosten ten bedrage van 1.000 NF gevraagd voor een volledige stage.

Inschrijvingen worden ingewacht op het Secretariaat van het *Bureau Universitaire de Recherche Opérationnelle*, 17 rue Richer, Paris 9e.

PROGRAMMA

EERSTE WEEK

1. — Grammatica der eindige verzamelingen — Algebra van Boole — Netwerken.
2. — Verbanden — Overeenkomsten — Functies.
3. — Vectorruimten — Lineaire rekentechnieken.
4. — Enkele combinatieproblemen en hun algorithmische voorstelling.
5. — Matrices en linaire vormen — Oplossing der vergelijkingen.
6. — Spectraalanalyse (opeenvolgende machten van een vierkante matrix).
7. — Lineaire operatoren : Stieltjes-integraal.
8. — Waarschijnlijkheidsleer — Begrip : toevalsveranderlijke.
   — Stochostisch verband — Het optellen van toevalsveranderlijken.
   — Herhaalde proefnemingen.
9. — Lineaire processen.
10. — Lineaire functievergelijkingen.
11. — De wet van de grote getallen.
12. — Entropie.
13. — Maxima- en minimaproblemen — Dualiteit — Ongelijkheden.
14. — Ergodiciteit van een lineair proces.
15. — Functieruimten : genererende en karakteristieke functies.
16. — Voornaamste wetten der waarschijnlijkheidsleer.
17. — Multiplicatoren van Lagrange — Stellingen van Kuhn en Tucker.
18. — Toevalsprocessen volgens Poisson.
19. — Toevalsprocessen volgens Markoff.
20. — Algemene bespreking.

21. — Economische rekentechnieken.
22. — De wiskunde der economische programmatie.
23. — Het verwekken van toevalsverschijnselen (toevalscijfers, Monte-Carlo methodes).
24. — Veelvlakken en lineaire programmatie.
25. — Beslissingsvraagstukken tegenover onzekerheid — De begrippen nut en waarschijnlijkheid.
26. — Dualiteit en lineaire programmatie.
27. — Statistische beslissingen — Klassieke theorie der toetsen — Theorie van Wald.
28. — Oplossing van lineaire programma's.
29. — Een dynamisch programma : oplossing door recurrentie en oplossing met de Simplex-methode.
30. — Wachttijdproblemen : het economisch beheer van de dode tijden.
31. — Enkele vraagstukken over Stockbeheer.
32. — Voorbeeld van toepassing op een speciaal geval.
33. — Voorbeeld van toepassing op een speciaal geval.
34. — Theorie der spelen : 1. Wiskundige theorie van het duel.
35. — Theorie der spelen : 2. De samenwerking en de strijd.
36. — Voorbeeld van toepassing op een speciaal geval.
37. — Enkele grondbeginselen van de theorie der dynamische programmatie.
38. — Een voorbeeld van rationalisatie der sekwente beslissingen over de prijsbepaling op een welbepaalde markt.
39. — Theorie der spelen : 3. Collectieve beslissingen en hun logische moeilijkheden (Condorceteffect, stelling van Arrow).
40. — Algemene bespreking.

Het merendeel dezer voordrachten wordt gehouden door de leden van het Bureau de Recherche Opérationnelle : HH. BOUZITAT, GIRAULT, GUILBAUD, KREWERAS, MORLAT, VALETTE.

Wij vernemen met genoegen dat de Heuschlingprijs toegekend werd aan de Heer F. BULTOT, lid van de « Société belge de Statistique », hoofd van het Klimatologisch Bureau van het I.N.E.A.C. en lid van de « Académie royale des Sciences d'Outre-Mer ».

Deze prijs bekroont het beste werk van een Belgisch auteur die een belangrijke bijdrage tot de statistiek levert. Hij wordt om de vijf jaar toegekend door een jury van negen leden die door de Koning worden gekozen onder de kandidaten voorgedragen door de Hoge Raad voor de Statistiek, de « Académie royale de Belgique » en de Koninklijke Vlammse Akademie voor Wetenschappen, Letteren en Schone Kunsten van België ».

Tussen de jongste publikaties van de Uitgeverij Eyrolles (61, bd St-Germain — Paris V°) ontdekten wij het boek :

LE CONTROLE STATISTIQUE DES FABRICATIONS

door René CAVE,

Ancien élève de l'Ecole polytechnique - Ingénieur militaire en chef de l'Armement.
Met een voorwoord door G. DARMOIS, *Membre de l'Institut.*
Tweede geheel herziene uitgave.

Dit boek beoogt een praktisch doel. Het geeft eenvoudige toepassingsmethoden voor statistische controle in de werkplaats. Op deze wijze zal het controlepersoneel

gemakkelijker het belang en het gebruik van deze nieuwe technieken begrijpen, zonder op de theorie te moeten ingaan.

Dank zij de volledige herwerking van het boek kon het aangevuld worden met een belangrijk aantal nieuwigheden terwijl de notaties genormaliseerd werden.

In het eerste deel vindt de lezer het nodige statistisch materiaal nl. de definities der fundamentele wetten, maar de schrijver heeft er tevens een stel praktische en nuttige beschouwingen aan toegevoegd zoals het interpreteren van frekwentiepolygonen, het gebruik van toevalscijfers, het verlies aan informatie, enz.

Het tweede deel behandelt de statistische methodes voor controle tijdens fabricage door meting en door vergelijking met kalibers (klassieke methodes, methodes uitgewerkt door de schrijver, recente nieuwe en vereenvoudigde methodes). Het werk van de controleur wordt vergemakkelijkt door tabellen en grafieken.

Het derde deel bevat de nodige aanwijzingen om een efficiënte afnamecontrole uit te werken, bij minimale of vastgestelde kostprijs, voor overeengekomen risicos van leverancier en afnemer of voor een bepaalde gemiddelde kwaliteitslimiet.

Het vierde deel heeft betrekking op de toepassing van statistische methodes op het onderzoek ; het bevat de onmisbare voorschriften tot het nemen van een beslissing met geminimaliseerde of op zijn minst vastgestelde risico's op vergissingen (vergelijken van apparaten, van fabricageprocédé's enz.).

Het laatste hoofdstuk omschrijft de werkzaamheden van een kwaliteitsdienst.

18 bijlagen (bewijsvoeringen), samenvattende tabellen (18 blz.), 20 tabellen (waaronder 4 nieuwe) en 9 grafieken vervolledigen het boek.

Op deze wijze herwerkt en aangevuld richt het boek van M. Cavé zich niet uitsluitend tot de statistici uit de industrie, maar tevens tot de ingenieurs en techniekers die belang hebben bij het gebruik van deze nieuwe methodes voor de controle van de fabricage.

PRIX DE VENTE

Au numéro	: Belgique	75	FB
	Etranger	90	FB
Abonnement	: Belgique	250	FB
(4 numéros)	Etranger	300	FB

TARIF DE PUBLICITE (4 numéros)

La page	:	5.000 F
La 1/2 page	:	3.000 F
Le 1/4 page	:	2.000 F

Les frais de clichés sont à charge de l'annonceur.

PUBLICATIONS D'ARTICLES

1) La Revue est ouverte aux articles traitant de statistique pure et appliquée, de Recherche Opérationnelle et de « Quality Control ».
2) Les manuscrits seront dactylographiés et peuvent être envoyés au secrétariat de la Revue : 10, rue du Tulipier à Bruxelles 19.
3) Les auteurs d'articles techniques recevront 20 tirés à part de leurs textes.
4) La responsabilité des articles n'incombe qu'à leurs auteurs.

VERKOOPPRIJS

Per nummer	: België	75	BF
	Buitenland	90	BF
Abonnement	: België	250	BF
(4 nummers)	Buitenland	300	BF

ADVERTENTIETARIEF (4 nummers)

Per bladzijde	:	5.000 F
Per 1/2 bladzijde	:	3.000 F
Per 1/4 bladzijde	:	2.000 F

De cliché-onkosten vallen ten laste van de adverteerders.

PUBLICATIES VAN ARTIKELS

1) Het tijdschrift neemt artikels aan over wiskundige statistiek en toepassingen, over operationeel onderzoek en kwaliteiszorg.
2) De teksten dienen getypt gestuurd te worden naar het secretariaat van het Tijdschrift : 10, Tulpenboomstraat, Brussel 19.
3) De auteurs ontvangen 20 overdrukken van de technische artikels.
4) De auteurs zijn alleen verantwoordelijk voor de inhoud van hun teksten.