OBSERVATIONS ON ANALYSIS OF VARIANCE THEORY

By Hilda Geiringer[1]

Bryn Mawr College

One of the important problems of theoretical statistics is the following. Let $x_1, x_2, \cdots x_N$ be the results of N observations; by means of these results we want to test the hypothesis that $V_i(x)$ is the distribution of the ith chance variable x_i. In that situation we often decide to choose a test function $F(x_1, x_2, \cdots x_N)$ and to determine the distribution of F under the above assumption. By means of this distribution we compute the probability of $\xi_1 \leqq F \leqq \xi_2$ and compare this result with the observed value of F.

Suppose there are m groups, each of n observations on $m \cdot n$ chance variables $x_{\mu\nu}$. We may test hypotheses regarding the mn distributions of the $x_{\mu\nu}$ in the way just mentioned. In analysis of variance theory we often use as test functions certain quadratic forms s_w^2 and s_a^2 ("variance within" and "among classes") and their quotient (multiplied by $m(n-1)/(m-1)$), usually denoted by z. Its distribution has been investigated by R. A. Fisher [2] under the assumption that the chance variables are mutually independent and subject to the same normal law. "The five per cent and one per cent points of this distribution have been tabulated by R. A. Fisher and are used to test, whether these two estimates of the same magnitude are significantly different. One gets thus a test of significance *to test whether our sample is a random sample from a homogeneous normal population or not.*[2] If the probability of a certain z-value is too small we shall reject the hypothesis that the sample is a random sample from a homogeneous normal population" [5].

The use of Fisher's z-test is also recommended if we may reasonably assume that the theoretical distributions are approximately normal. "Unless some rather startling lack of normality is known or suspected analysis of variance may be used with confidence." This last remark can be understood by considering that, as we will see in detail, some of the basic results of our theory are independent of the normality of the populations. It is however this assumption of normality which makes possible the complete and elegant solution of the problem of distribution obtained by R. A. Fisher.

If it is *not* possible to determine the exact distribution of a test function under sufficiently general assumptions we may:

a) make simple and particular assumptions concerning the populations

b) investigate an asymptotic solution of the problem, i.e. determine the distributions of the test functions for large samples,[3] or

c) study the mathematical expectations and the variances of the test functions

[1] Research under a grant in aid of the American Philosophical Society.

[2] My italics.

[3] cf. statement (a) page 355.

Institute of Mathematical Statistics is collaborating with JSTOR to digitize, preserve, and extend access to *The Annals of Mathematical Statistics.* www.jstor.org

for small samples under appropriately general assumptions regarding the populations (this should be done independently of concepts of estimation, unbiased estimate etc.).

This last procedure provides us with tests which suffice in actual practice.[4]

It is well known that the expectations of the two forms $s_a^2/(m-1)$ and $s_w^2/m(n-1)$ are the same even if the populations are not normal, but equal each other (*Bernoulli* series). In addition we shall prove the theorem, familiar in case of the Lexis quotient [9], that under these conditions the *expectation of their quotients equals unity* (section 1, (b)). The next step consists in investigating certain inequalities characteristic of *Lexis* or *Poisson* series. The different criteria will be completed by the computation of the respective variances (Section 1, (c)).

In addition to the above mentioned test functions other symmetrical test functions have been considered [5]. In studying these we shall again assume general populations. It will be seen that the Lexis as well as the Poisson series may be characterized by equalities (instead of inequalities) (Section 2, (a)), and we can generalize our theorem on the expectation of the quotient (Section 2, (b)) to this case. Then the variances of these test functions will be investigated.

It seems worthwhile *to omit the assumption of independence* of the chance variables and to study different kinds of mutual dependence. These investigations lead to interesting relations among the expectations[5] (Section 2, (c)). They seem to be related to Fisher's "intraclass correlation" and to supplement his idea.

Most of the results of Sections 1 and 2 can be generalized to the analysis of covariance (section 3).

1. Variance within and among classes.

(*a*). *The test functions.* Let $x_{\mu\nu}$ ($\mu = 1, \cdots m$; $\nu = 1, \cdots n$) be $m \cdot n$ chance variables and put

$$(1)\qquad \begin{aligned} a_\mu &= \frac{1}{n}\sum_{\nu=1}^{n} x_{\mu\nu}, \qquad \bar{a}_\nu = \frac{1}{m}\sum_{\mu=1}^{m} x_{\mu\nu}, \\ a &= \frac{1}{mn}\sum_{\mu=1}^{m}\sum_{\nu=1}^{n} x_{\mu\nu} = \frac{1}{m}\sum_{\mu=1}^{m} a_\mu = \frac{1}{n}\sum_{\nu=1}^{n} \bar{a}_\nu . \end{aligned}$$

[4] The important paper of Irwin [5] assumes normality of the populations. H. L. Rietz [8] computes the expectations of s_a^2 and s_w^2 under rather general assumptions for the populations and considers the cases of Bernoulli, Lexis, Poisson series, but does not consider tests of significance; nor does he consider the symmetric test functions (section 2 of this paper). In later papers on our subject the assumption of normal and independent populations recurs. Another approach [11] in the problem of analysis of variance is to use ranks instead of the actual values (this has been pointed out by the referee to the author, who is very grateful for this comment).

[5] They generalize previous results of the author.

We then introduce the three quadratic forms

$$(2)\quad s^2 = \sum_\mu \sum_\nu (x_{\mu\nu} - a)^2; \qquad s_a^2 = n \sum_\mu (a_\mu - a)^2; \qquad s_w^2 = \sum_\mu \sum_\nu (x_{\mu\nu} - a_\mu)^2,$$

with the respective ranks (degrees of freedom)

$$(3)\qquad r = mn - 1, \qquad r_a = m - 1, \qquad r_w = m(n - 1).$$

Then we have

$$(4)\qquad s^2 = s_a^2 + s_w^2, \qquad r = r_a + r_w.$$

The $m \cdot n$ theoretical distributions are assumed in this section to be independent of each other. Let $V_{\mu\nu}(x)$ be the probability that $x_{\mu\nu} \leqq x$ and

$$(5)\qquad \alpha_{\mu\nu} = \int x\, dV_{\mu\nu}(x), \qquad \sigma_{\mu\nu}^2 = \int (x - \alpha_{\mu\nu})^2\, dV_{\mu\nu}(x),$$

where the integrals are *Stieltjes* integrals; thus the $V_{\mu\nu}(x)$ may be e.g. general arithmetical or geometrical distributions.[6]

Let us compute the mathematical expectation of the three test functions with respect to the $m \cdot n$-dimensional distribution:

$$V_{11}(x_{11}) V_{12}(x_{12}) \cdots V_{mn}(x_{mn}).$$

$$(6)\quad E[F(x_{11}, \cdots x_{mn})] = \int \cdots \int F(x_{11}, \cdots x_{mn})\, dV_{11}(x_{11}) \cdots dV_{mn}(x_{mn}).$$

We have then

$$(7)\qquad E\left[\frac{s^2}{mn - 1}\right] = \frac{1}{mn} \Sigma\Sigma\sigma_{\mu\nu}^2 + \frac{1}{mn - 1} \Sigma\Sigma(\alpha_{\mu\nu} - \alpha)^2,$$

$$(8)\qquad E\left[\frac{s_{a^2}}{m - 1}\right] = \frac{1}{mn} \Sigma\Sigma\sigma_{\mu\nu}^2 + \frac{1}{m - 1} \cdot n\Sigma(\alpha_\mu - \alpha)^2,$$

$$(9)\qquad E\left[\frac{s_w^2}{m(n - 1)}\right] = \frac{1}{mn} \Sigma\Sigma\sigma_{\mu\nu}^2 + \frac{1}{m(n - 1)} \Sigma\Sigma(\alpha_{\mu\nu} - \alpha_\mu)^2.$$

From these equalities we deduce:

1. *If the $m \cdot n$ theoretical mean values $\alpha_{\mu\nu}$ are all equal* (Bernoulli series), *then the expectations in* (6), (7), (8) *are equal; i.e.*

$$(10)\qquad E_B\left(\frac{s^2}{mn - 1}\right) = E_B\left(\frac{s_a^2}{m - 1}\right) = E_B\left(\frac{s_w^2}{m(n - 1)}\right).$$

2. *If the $\alpha_{\mu\nu}$ are equal "by rows" but differ from row to row* (Lexis series), i.e. $\alpha_{\mu\nu} = \alpha_\mu$ *but* $\alpha_\mu \neq \alpha$. *Then*

---

[6] $V_{\mu\nu}(x)$ is a monotone non-decreasing function. Hence it has at most a denumerable set of ordinary jump discontinuities; at such a point it is continuous to the right but not to the left. Moreover it possesses a finite derivative $v_{\mu\nu}(x)$ almost everywhere.

$$(11) \qquad E_L\left[\frac{s_a^2}{m-1} - \frac{s^2}{mn-1}\right] = \frac{mn(n-1)}{(m-1)(mn-1)} \sum_\mu (\alpha_\mu - \alpha)^2 > 0,$$

$$(12) \qquad E_L\left[\frac{s_a^2}{m-1} - \frac{s_w^2}{m(n-1)}\right] = \frac{n}{m-1} \sum_\mu (\alpha_\mu - \alpha)^2 > 0.$$

3. *If the $\alpha_{\mu\nu}$ are equal "by columns" but differ from column to column* (Poisson series), *then $\alpha_{\mu\nu} = \bar{\alpha}_\nu$; $\alpha_\mu = \alpha$ and*

$$(13) \qquad E_P\left[\frac{s_a^2}{m-1} - \frac{s^2}{mn-1}\right] = -\frac{m}{mn-1} \sum_\nu (\bar{\alpha}_\nu - \alpha)^2 < 0,$$

$$(14) \qquad E_P\left[\frac{s_a^2}{m-1} - \frac{s_w^2}{m(n-1)}\right] = -\frac{1}{n-1} \sum_\nu (\bar{\alpha}_\nu - \alpha)^2 < 0.$$

In the Lexis theory[7] we speak of *normal, supernormal* or *subnormal* dispersion depending on whether the *observed* value of $\frac{s_a^2}{m-1}$ is *equal, greater* or *less* than that of $\frac{s^2}{mn-1}$ and we usually consider the quotient

$$(15) \qquad L = \frac{s_a^2}{m-1} \Big/ \frac{s^2}{mn-1}.$$

In analysis of variance theory we usually compare $s_a^2/(m-1)$ (variance among rows) with $s_w^2/m(n-1)$ (variance within rows) and introduce the quotient

$$(16) \qquad V = \frac{s_a^2}{m-1} \Big/ \frac{s_w^2}{m(n-1)}.$$

It follows from (4): *If $L \gtreqless 1$ then $V \gtreqless 1$ and conversely.* We may therefore speak of *normal* or *non-normal dispersion* with respect either to L or to V.

The results given by equations (10)–(14) can be expressed as follows: *If the $m \cdot n$ theoretical distributions are all equal the mathematical expectation of s^2/r, of s_a^2/r_a and of s_w^2/r_w are equal. In the case of a Lexis series the expectation of s_a^2/r_a is greater than s^2/r and greater than s_w^2/r_w and in the case of a Poisson series the opposite is true.*

We generally use these facts in order to *make inferences about the unknown populations from the observed values of our test functions $V_{\mu\nu}(x)$.* If e.g., the observed value of s_a^2/r_a is "significantly"[8] greater than that of s^2/r we may assume that the theoretical distributions form a Lexis series. But of course such a significant deviation can also be explained by quite different assumptions regarding the populations (see Section 2, (c)).

(*b*). *Mathematical expectation of the quotient of the test functions.* We are going to prove in this section a theorem of some mathematical interest. This theorem is a generalization of an analogous theorem in the Lexis theory [9].

[7] The relation between these considerations and the Lexis theory will be dealt with in another paper.

[8] The meaning of the word "significantly" has of course still to be explained.

We have seen (10) that the mathematical expectations, defined by (6), of the three test functions

$$S = \frac{s^2}{mn - 1}, \qquad S' = \frac{s_a^2}{m - 1}, \qquad S'' = \frac{s_w^2}{m(n - 1)},$$

are equal *if the $m \cdot n$ populations are equal* (*i.e. have identical distributions*). *We will show that even in this case*

$$E\left(\frac{S'}{S}\right) = 1, \qquad E\left(\frac{S''}{S}\right) = 1. \tag{17}$$

Let us put $m \cdot n = N$, and let the N chance variables be arranged in a one-dimensional sequence. As S' and S are of second degree in the x_ν ($\nu = 1, 2, \cdots N$) we may write

$$S' - S = A + \sum B_\nu x_\nu + \sum C_\nu x_\nu^2 + \sum_{\nu \neq \rho} D_{\nu\rho} x_\nu x_\rho$$

where the A, B_ν, C_ν and $D_{\nu\rho}$ are constants. Now form the expectation, defined by (6), of $(S' - S)$ under the assumption that the N populations are equal $V_\nu(x) = V(x)$ ($\nu = 1 \cdots N$). Denoting by α and σ^2 the mean value and variance of $V(x)$ and putting $\Sigma B_\nu = B$, $\Sigma C_\nu = C$, $\Sigma D_{\nu\rho} = D$ we find

$$E(S' - S) = A + B\alpha + C(\sigma^2 + \alpha^2) + D\alpha^2 = 0.$$

And as this equality holds for an arbitrary distribution $V(x)$, we deduce that $A = B = C = D = 0$. Let us then compute under the same assumption the expectation of $(S' - S)/S$. Now the expectations of $1/S$, x_ν/S, x_ν^2/S, $x_\nu x_\rho/S$, take the place of the expectations of 1, x_ν, x_ν^2, $x_\nu x_\rho$. But these new expectations are also *independent of the index ν*, because of the *equality of the N populations* and the *symmetry* of S in the N variables $x_1, \cdots x_N$. Hence we may write

$$E\left(\frac{1}{S}\right) = \mu_0, \qquad E\left(\frac{x_\nu}{S}\right) = \mu_1, \qquad E\left(\frac{x_\nu^2}{S}\right) = \mu_2, \qquad E\left(\frac{x_\nu x_\rho}{S}\right) = \mu_3,$$

and we find

$$E\left(\frac{S' - S}{S}\right) = E\left(\frac{S'}{S} - 1\right) = A\mu_0 + B\mu_1 + C\mu_2 + D\mu_3 = 0,$$

because $A = B = C = D = 0$. Hence $E(S'/S) = 1$.

We may prove in the same way that $E(S''/S) = 1$.

We have however proved (17) only under the assumption that all the N *populations are equal*, whereas (10) is true under the mere hypothesis that the mean values of the populations $V_\nu(x)$ are the same.

(*c*). *The variances of the test functions*. The distribution of our test functions and of their quotients V or L have been determined and tabulated by R. A. Fisher under the hypothesis that the $m \cdot n$ chance variables are *independent* and obey the same normal Gaussian law. Consequently by means of Fisher's distribution we can test only the hypothesis that the theoretical populations have *both these* properties.

If in a statistical problem it is not possible to determine the exact distributions of the test functions under sufficiently general assumptions regarding the populations, one of the following procedures is frequently used:
a) one tries to find an *asymptotic* solution of the problem, i.e. to determine the distribution of the test functions in question for *large samples*. The distribution of the analysis of variance quotient, as n tends to infinity, has been established by W. G. Madow [6]. The same problem for the Lexis quotient was solved as early as 1873 by Helmert [4]. *As m tends to infinity* the limiting distribution is a Gaussian distribution, which follows from general theorems of v. Mises [7].
b) For *small* samples, i.e. if m and n are finite we may determine the expectations and the variances of the test functions for appropriately general populations and establish in this way a test of significance.

In this section we shall compute the variances of our test functions. Let us first assume arbitrary but equal populations $V_\nu(x) = V(x)$ and denote by M_i the ith moment about the mean of $V(x)$:

$$(18)\qquad \mathrm{M}_i = \int (x-\alpha)^i\, dV(x), \qquad (i = 1, 2, \cdots), \qquad \alpha = \int x\, dV(x), \qquad \mathrm{M}_2 = \sigma^2.$$

Then we find immediately the variance of $S = \dfrac{s^2}{mn-1}$ using a well-known formula for the variance of a sample variance

$$(19)\qquad \mathrm{Var}\left\{\frac{s^2}{mn-1}\right\} = \mathrm{Var}\left\{\frac{\Sigma\Sigma(x_{\mu\nu}-a)^2}{mn-1}\right\} = \frac{1}{mn}\left\{\mathrm{M}_4 - \frac{mn-3}{mn-1}\mathrm{M}_2^2\right\}.$$

If we need the analogous variance in case of different populations we let

$$t^2 = \sum_{\rho=1}^{r}(y_\rho - b)^2 \qquad \text{where } b = \frac{1}{r}(y_1 + \cdots + y_r)$$

and let $V_\rho(y)$, $(\rho = 1, \cdots r)$, be r populations, and

$$\beta_\rho = \int y\, dV_\rho(y), \qquad \frac{1}{r}\sum_{\rho=1}^{r}\beta_\rho = \beta,$$

$$\int (y-\beta_\rho)^i\, dV_\rho(y) = \mu_i^{(\rho)}, \qquad (i = 1, 2, \cdots, \rho = 1, 2, \cdots r), \mu_2^{(\rho)} = \sigma_\rho^2.$$

Then the following formula may be used:

$$(20)\qquad \mathrm{Var}(t^2) = \left(\frac{r-1}{r}\right)^2 \sum_{\rho=1}^{r}[\mu_4^{(\rho)} - \sigma_\rho^4] + 4\,\frac{r-1}{r}\sum_{\rho=1}^{r}\mu_3^{(\rho)}(\beta_\rho-\beta)^2 + 4\sum_{\rho=1}^{r}\sigma_\rho^2(\beta_\rho-\beta)^2 + \frac{4}{r^2}\sum_{\rho<\tau}\sigma_\rho^2\sigma_\tau^2.$$

We may check (20) by putting the $V_\rho(y)$ all equal to $V(y)$ and find

$$(20') \qquad \operatorname{Var}(t^2) = \frac{r-1}{r}[(r-1)\mu_4 - (r-3)\sigma^4],$$

in accordance with (19).

In order to determine the variance of s_a^2 by means of these formulae we consider $\frac{1}{m}\sum_\mu (a_\mu - a)^2$ as a sample variance. The n distributions in the nth row are $V_{\mu1}(x), V_{\mu2}(x), \cdots V_{\mu n}(x)$. Or, if we assume that they are all equal, simply $V(x) = V(x_{\mu\nu})$. Let us put $\frac{1}{n}x_{\mu\nu} = z_{\mu\nu}$ and $V(x_{\mu\nu}) = V'(z_{\mu\nu})$, and denote by $W(a_\mu)$ the distribution of the average of the elements in the μth row:

$$W(a_\mu) = \int\cdots\int dV'(z_{\mu1})\,dV'(z_{\mu2})\cdots dV'(z_{\mu,n-1})V'(a_\mu - z_{\mu1} - \cdots - z_{\mu,n-1}).$$

There is such a distribution for each row, and we have to find the variance of $\sum_\mu (a_\mu - a)^2$ with respect to the combination of these m distributions. In order to be able to apply (20') we need the second and fourth moments of these distributions. We have for the mean value α' of $W(a_\mu)$:

$$\alpha' = n\cdot(\text{mean value of } V') = n\cdot\frac{1}{n}\alpha_\mu = \alpha$$

and for the variance μ_2' of $W(a_\mu)$: $\mu_2' = \frac{\sigma^2}{n}$. We still need μ_4'. By repeated use of the formula

$$\iint [(x_1 - a_1) + (x_2 - a_2)]^4\,dV(x_1)\,dV(x_2)$$
$$= \int (x_1 - a_1)^4\,dV(x_1) + \int (x_2 - a_2)^4\,dV(x_2)$$
$$+ 6\int (x_1 - a_1)^2\,dV(x_1)\int (x_2 - a_2)^2\,dV(x_2),$$

and of the fact that $W(a_\mu)$ is simply the distribution of the sum of n variables $z_{\mu\nu}$ we get:

$$\mu_4' = \frac{1}{n^4}\left(n\mathrm{M}_4 + 6\,\frac{n(n-1)}{2}\,\mathrm{M}_2^2\right) = \frac{1}{n^3}(\mathrm{M}_4 + 3(n-1)\mathrm{M}_2^2)$$

where M_4 and M_2 are the values introduced in (18).

We now apply (20') and get

$$\operatorname{Var}[\Sigma(a_\mu - a)^2] = \frac{m-1}{m}[(m-1)\mu_4' - (m-3)\mu_2'].$$

and substituting the values of μ_2' and μ_4', we find by an easy computation the final result:

$$(21)\qquad \operatorname{Var}\left\{\frac{n}{m-1}\,\Sigma(a_\mu - a)^2\right\} = \frac{1}{mn}\,(M_4 - 3M_2^2) + \frac{2}{m-1}\,M_2^2 .$$

If we compare this last formula with (19) we see that the right side in (21) is of order $1/m$, whereas that in (19) is of order $1/mn$. Therefore, for sufficiently large values of n, s^2/r will be "more exact" than s_a^2/r_a. In some presentations of the Lexis theory it is implied that the value s_a^2/r_a is to be compared with the theoretical or exact value s^2/r; we may see a certain justification for this idea in the result just mentioned. This may lead us also to use s^2/r as an unbiased estimate of the unknown population variance if $\alpha_{\mu\nu} = \alpha$ (see (7) and (8)).

By means of the simple formulae (19) and (21) we can now easily test whether the values of s^2/r and s_a^2/r_a whose expectations are equal in case of equal populations differ significantly from each other. Of course we must compute as usual approximate values of M_2 and M_4 from the observations. If n is comparatively large—as it usually is e.g. in the Lexis theory—only the term $\frac{2}{m-1}\,M_2^2$ will be significant. If the hypothetical population is Gaussian ($M_4 = 3M_2^2$) the right side of (21) reduces to $\frac{2}{m-1}\,M_2^2$ and that of (19) to $\frac{2M_2^2}{mn-1}$; hence these variances are in the ratio of $\frac{1}{r_a}\Big/\frac{1}{r}$, as one might expect.

2. Symmetric Test Functions.

(*a*). *New equalities for Lexis and Poisson series.* In Section 1, starting with the formula $s^2 = s_a^2 + s_w^2$ we used the test functions s^2/r, s_a^2/r_a, s_w^2/r_w. This implied a difference between rows and columns, which is often justified, e.g. in the Lexis theory. The following decomposition of s^2 is symmetric with respect to rows and columns. Let

$$(1)\qquad \begin{gathered} \frac{1}{n}\sum_{\nu=1}^{n} x_{\mu\nu} = a_\mu, \qquad \frac{1}{m}\sum_{\mu=1}^{m} x_{\mu\nu} = \bar{a}_\nu, \\ \frac{1}{mn}\sum_{\mu=1}^{m}\sum_{\nu=1}^{n} x_{\mu\nu} = \frac{1}{m}\sum_{\mu=1}^{m} a_\mu = \frac{1}{n}\sum_{\nu=1}^{n} \bar{a}_\nu = a, \end{gathered}$$

and

$$(2)\qquad \begin{aligned} &s^2 = \Sigma\Sigma(x_{\mu\nu} - a)^2, \qquad s_a^2 = n\Sigma(a_\mu - a)^2, \qquad s_w^2 = \Sigma\Sigma(x_{\mu\nu} - a_\mu)^2 \\ &S^2 = \Sigma\Sigma(x_{\mu\nu} - a_\mu - \bar{a}_\nu + a)^2, \qquad \bar{s}_a^2 = m\Sigma(\bar{a}_\nu - a)^2, \qquad \bar{s}_w^2 = \Sigma\Sigma(x_{\mu\nu} - \bar{a}_\nu)^2 \end{aligned}$$

with the respective ranks

$$(3)\qquad \begin{aligned} &r = mn - 1, \qquad r_a = m - 1, \qquad r_w = m(n-1), \\ &R = (m-1)(n-1), \qquad \bar{r}_a = n - 1, \qquad \bar{r}_w = n(m-1). \end{aligned}$$

Then

$$(5)\qquad s^2 = s_a^2 + \bar{s}_a^2 + S^2 = s_a^2 + s_w^2 = \bar{s}_a^2 + \bar{s}_w^2$$

and

$$r = r_a + \bar{r}_a + R = r_a + r_w = \bar{r}_a + \bar{r}_w . \tag{6}$$

We find the expectations of these forms under the assumptions, of arbitrary populations $V_{\mu\nu}(x)$ which are independent and different from each other. We then specialize for Bernouilli series, Lexis and Poisson series of populations respectively. Denoting by $\alpha_{\mu\nu}$ and $\sigma^2_{\mu\nu}$ the mean value and variance of $V_{\mu\nu}(x)$ and by

$$\alpha_\mu = \frac{1}{n}\sum_\nu \alpha_{\mu\nu}, \qquad \bar{\alpha}_\nu = \frac{1}{m}\sum_\mu \alpha_{\mu\nu}, \qquad \alpha = \frac{1}{m}\sum \alpha_\mu = \frac{1}{n}\sum \bar{\alpha}_\nu, \tag{6}$$

we find for the expected values defined in (6) Section 1:

$$\begin{aligned}
E\left[\frac{s^2}{mn-1}\right] &= \frac{1}{mn}\Sigma\Sigma\sigma^2_{\mu\nu} + \frac{1}{mn-1}\Sigma\Sigma(\alpha_{\mu\nu}-\alpha)^2,\\
E\left[\frac{s_a^2}{m-1}\right] &= \frac{1}{mn}\Sigma\Sigma\sigma^2_{\mu\nu} + \frac{1}{m-1}n\Sigma(\alpha_\mu-\alpha)^2,\\
E\left[\frac{\bar{s}_a^2}{n-1}\right] &= \frac{1}{mn}\Sigma\Sigma\sigma^2_{\mu\nu} + \frac{1}{n-1}m\Sigma(\bar{\alpha}_\nu-\alpha)^2,\\
E\left[\frac{S^2}{(m-1)(n-1)}\right] &= \frac{1}{mn}\Sigma\Sigma\sigma^2_{\mu\nu} + \frac{1}{(m-1)(n-1)}\Sigma\Sigma(\alpha_{\mu\nu}-\alpha_\mu-\bar{\alpha}_\nu+\alpha)^2,\\
E\left[\frac{s_w^2}{m(n-1)}\right] &= \frac{1}{mn}\Sigma\Sigma\sigma^2_{\mu\nu} + \frac{1}{m(n-1)}\Sigma\Sigma(\alpha_{\mu\nu}-\alpha_\mu)^2,\\
E\left[\frac{\bar{s}_w^2}{n(m-1)}\right] &= \frac{1}{mn}\Sigma\Sigma\sigma^2_{\mu\nu} + \frac{1}{n(m-1)}\Sigma\Sigma(\alpha_{\mu\nu}-\bar{\alpha}_\nu)^2.
\end{aligned} \tag{7}$$

In the Bernouilli case which as far as the author knows is the only one which has been considered in this connection [5], we get the wellknown result:

$$\begin{aligned}
E_B\left[\frac{s^2}{mn-1}\right] &= E_B\left[\frac{s_a^2}{m-1}\right] = E_B\left[\frac{\bar{s}_a^2}{n-1}\right]\\
&= E_B\left[\frac{s_w^2}{m(n-1)}\right] = E_B\left[\frac{\bar{s}_w^2}{n(m-1)}\right] = E_B\left[\frac{S^2}{(m-1)(n-1)}\right].
\end{aligned} \tag{8}$$

Now let us assume a Lexis series, with

$$\alpha_{\mu\nu} = \alpha_\mu\,; \qquad \alpha_\mu \neq \alpha; \qquad \bar{\alpha}_\nu = \alpha, \qquad \sigma^2_{\mu\nu} = \sigma^2_\mu \tag{9}$$

Then (7) reduces to

$$\begin{aligned}
E_L\left[\frac{s^2}{mn-1}\right] &= \frac{1}{m}\Sigma\sigma^2_\mu + \frac{1}{mn-1}n\Sigma(\alpha_\mu-\alpha)^2,\\
E_L\left[\frac{s_a^2}{m-1}\right] &= \frac{1}{m}\Sigma\sigma^2_\mu + \frac{1}{m-1}\cdot n\Sigma(\alpha_\mu-\alpha)^2,\\
E_L\left[\frac{\bar{s}_a^2}{n-1}\right] &= \frac{1}{m}\Sigma\sigma^2_\mu,
\end{aligned}$$

$$E_L\left[\frac{S^2}{(m-1)(n-1)}\right] = \frac{1}{m}\,\Sigma\sigma_\mu^2,$$
$$E_L\left[\frac{s_w^2}{m(n-1)}\right] = \frac{1}{m}\,\Sigma\sigma_\mu^2,$$
$$E_L\left[\frac{\bar{s}_w^2}{n(m-1)}\right] = \frac{1}{m}\,\Sigma\sigma_\mu^2 + \frac{1}{m-1}\,\Sigma(\alpha_\mu - \alpha)^2.$$

From these formulae we deduce—besides the inequalities (11), (12) of Section 1, and the corresponding formulae where the role of rows and columns is interchanged—the further inequalities:

$$(11)\qquad E_L\left[\frac{s_a^2}{m-1}\right] > E_L\left[\frac{\bar{s}_w^2}{n(m-1)}\right] > E_L\left[\frac{\bar{s}_a^2}{n-1}\right].$$

But there are also characteristic equalities, namely:

$$(12)\qquad E_L\left[\frac{\bar{s}_a^2}{n-1}\right] = E_L\left[\frac{S^2}{(m-1)(n-1)}\right] = E_L\left[\frac{s_w^2}{m(n-1)}\right].$$

These *equalities*[9] seem often to be more appropriate than the usual *inequalities* in testing the hypothesis of a Lexis series.

Let us finally consider the Poisson case which is very often neglected. There we have:

$$(13)\qquad \alpha_{\mu\nu} = \bar{\alpha}_\nu, \qquad \bar{\alpha}_\nu \neq \alpha, \qquad \alpha_\mu = \alpha, \qquad \sigma_{\mu\nu}^2 = \bar{\sigma}_\nu^2.$$

Then—beside the inequalities (13), (14) of Section 1 and the corresponding ones where the role of rows and columns is interchanged—we find the new *inequality*:

$$(14)\qquad E_P\left[\frac{s_a^2}{m-1} - \frac{\bar{s}_a^2}{n-1}\right] = -\frac{m}{n-1}\sum_\nu(\bar{\alpha}_\nu - \alpha)^2 < 0,$$

which of course corresponds to the Lexis inequality (11). The characteristic *equalities* are now:

$$(15)\qquad E_P\left[\frac{s_a^2}{m-1}\right] = E_P\left[\frac{S^2}{(m-1)(n-1)}\right] = E_P\left[\frac{\bar{s}_w^2}{n(m-1)}\right].$$

These equalities (12) and (15) can be used in testing the hypothesis of Lexis or Poisson series respectively in the same way as the equalities (9) for the Bernouilli case. We shall deal with the variances of these test functions in (d) of this section.

(*b*). *Mathematical expectations of the quotients of certain test functions.* We have seen that in case of a Lexis-Series the expectations of $\frac{\bar{s}_a^2}{n-1}$, of $\frac{S^2}{(m-1)(n-1)}$ and of $\frac{s_w^2}{m(n-1)}$ are equal. We will show that even in this case

[9] See [10] pp. 81–90 for proofs of these inequalities for the case of normal populations.

$$
(16)\qquad
\begin{aligned}
E_L\left[\frac{\bar{s}_a^2}{n-1}\Big/\frac{s_w^2}{m(n-1)}\right] &= 1,\\
E_L\left[\frac{s_w^2}{m(n-1)}\Big/\frac{S^2}{(m-1)(n-1)}\right] &= 1,\\
E_L\left[\frac{\bar{s}_a^2}{n-1}\Big/\frac{S^2}{(m-1)(n-1)}\right] &= 1,\\
E_L\left[\frac{S^2}{(m-1)(n-1)}\Big/\frac{s_w^2}{m(n-1)}\right] &= 1.
\end{aligned}
$$

Let us write for the moment: $\frac{\bar{s}_a^2}{n-1} = \bar{T}$ and $\frac{S^2}{(m-1)(n-1)} = T$. As both T and $\bar{T}$ are of second degree in the $x_{\mu\nu}$ we may write:

$$\bar{T} - T = A + \sum_{\mu,\nu} B_{\mu\nu} x_{\mu\nu} + \sum_{\mu,\nu} C_{\mu\nu} x_{\mu\nu}^2 + \sum\sum D_{\mu_1,ij;\,\mu_2,j}\, x_{\mu_1,i}\, x_{\mu_2 j},$$

where the A, B, C, D are constants. The last sum contains $\frac{1}{2}\cdot mn(mn - 1)$ terms and *not both* $\mu_1 = \mu_2$ and $i = j$ hold. Compute the expectation of $\bar{T} - T$ with respect to populations which form a Lexis series $V_{\mu\nu}(x) = V_\mu(x)$. Denote by α_μ, σ_μ^2 the respective mean values and variances. We then have because of (11):

$$0 = E_L[\bar{T} - T] = A + \sum_\mu \alpha_\mu \sum_\nu B_{\mu\nu} + \sum_\mu (\sigma_\mu^2 + \alpha_\mu^2) \sum_\nu C_{\mu\nu} + \sum_{\mu_1,\mu_2} \alpha_{\mu_1}\alpha_{\mu_2} \sum_{i,j} D_{\mu_1 i;\mu_2 j}$$

or introducing $\sum_\nu B_{\mu\nu} = B_\mu$; $\sum_\nu C_{\mu\nu} = C_\mu$; $\sum_{i,j} D_{\mu_1 i;\mu_2 j} = D_{\mu_1\mu_2}$ we get:

$$0 = E_L[\bar{T} - T] = A + \sum_\mu \alpha_\mu B_\mu + \sum_\mu (\sigma_\mu^2 + \alpha_\mu^2) C_\mu + \sum_{\mu_1,\mu_2} \alpha_{\mu_1}\alpha_{\mu_2} D_{\mu_1\mu_2}\,.$$

As this equality is exact for an arbitrary set of $V_\mu(x)$ we deduce that $A = 0$, $B_\mu = 0$, $C_\mu = 0$, $D_{\mu_1\mu_2} = 0$.

Let us now compute under the same assumption the expectation of $(\bar{T} - T)/T$. Here the expectations of $1/T$, $x_{\mu\nu}/T$ etc. will take the place of the expectations of 1, $x_{\mu\nu}$, $\cdots$. But these new expectations will not depend on the index ν (index within the row) because the populations are the same within each row and because of the symmetry of T in the $m\cdot n$ variables $x_{\mu\nu}$. Hence we can put

$$E\left(\frac{1}{T}\right) = l_0, \qquad E\left(\frac{x_{\mu\nu}}{T}\right) = l_\mu, \qquad E\left(\frac{x_{\mu\nu}^2}{T}\right) = \bar{l}_\mu, \qquad E\left(\frac{x_{\mu_1 i}\, x_{\mu_2 j}}{T}\right) = l_{\mu_1\mu_2}, \quad \text{etc.}$$

and we get

$$E\left[\frac{\bar{T} - T}{T}\right] = E\left(\frac{\bar{T}}{T} - 1\right) = Al_0 + \sum_\mu l_\mu B_\mu + \sum_\mu \bar{l}_\mu C_\mu + \sum_{\mu_1,\mu_2} l_{\mu_1,\mu_2} D_{\mu_1,\mu_2} = 0,$$

because all the coefficients are equal to zero. Our theorem is thus proved. The same conclusion holds if the denominator—without being symmetric in all the

$m \cdot n$ variables—does not depend on the row index. And as this last property holds for s_w^2 the expectations (16) are all shown to be equal to one.

Analogous relations are valid for Poisson series.

(c). *Non-independent populations.* We omit in this section the assumption of independence of the $m \cdot n$ populations but assume the theoretical population to be a general $m \cdot n$-variate distribution:

$$V(x_{11}, x_{12}, \cdots x_{mn}). \tag{17'}$$

From $V(x_{11}, x_{12}, \cdots x_{mn})$ we derive the mn one-dimensional distributions $V_{\mu\nu}(x)$ ($\mu = 1, \cdots m$; $\nu = 1, \cdots n$) by letting all the variables except $x_{\mu\nu}$ tend to $+\infty$, because $V_{\mu\nu}(x)$ is the probability that $x_{\mu\nu} \leqq x$ regardless of the values of the other variables. In a similar way we derive the $\frac{1}{2}mn(mn - 1)$ two dimensional distributions $V_{\mu_1\nu_1;\mu_2\nu_2}(x, y)$, that is the probability that $x_{\mu_1\nu_1} \leqq x$ and $x_{\mu_2\mu_2} \leqq y$. We get this distribution from (17′) as all the variables with the exception of $x_{\mu_1\nu_1}$ and $x_{\mu_2\nu_2}$ tend to $+\infty$. We denote as before by $\alpha_{\mu\nu}$ and $\sigma_{\mu\nu}^2$ the expectation of $x_{\mu\nu}$ and $(x_{\mu\nu} - \alpha_{\mu\nu})^2$ respectively. But the expectation of $(x_{\mu_1\nu_1} - \alpha_{\mu_1\nu_1})(x_{\mu_2\nu_2} - \alpha_{\mu_2\nu_2})$ which was zero in case of the independence of $x_{\mu_1\nu_1}$ and $x_{\mu_2\nu_2}$ may now differ from zero. Denote by $\mathfrak{E}$ the expectation with respect to (17′). Then:

$$\begin{aligned} &\mathfrak{E}[(x_{\mu_1\nu_1} - \alpha_{\mu_1\nu_1})(x_{\mu_2\nu_2} - \alpha_{\mu_2\nu_2})] \\ &= \iint \cdots \int (x_{\mu_1\nu_1} - \alpha_{\mu_1\nu_1})(x_{\mu_2\nu_2} - \alpha_{\mu_2\nu_2})\, dV(x_{11}, \cdots x_{mn}) \\ &= \iint (x - \alpha_{\mu_1\nu_1})(y - \alpha_{\mu_2\nu_2})\, dV_{\mu_1\nu_1;\nu_2\mu_2}(xy) = R_{\mu_1\nu_1;\mu_2\nu_2} = R_{\mu_2\nu_2;\mu_1\nu_1}. \end{aligned} \tag{17}$$

Let us first deduce a general formula for the expectation of a sample variance in the case of dependent populations. Let $P(y_1, \cdots y_r)$ be the distribution of r chance variables $y_1, \cdots y_r$ which have the average b. Denoting by β_ρ the expectation of y_ρ with respect to P, by β the average of the β_ρ, by τ_ρ^2 the expectation of $(y_\rho - \beta_\rho)^2$ by R_{ij} that of $(y_i - \beta_i)(y_j - \beta_j)$ we find, without difficulty, for the expectation of the sample variance

$$\begin{aligned} &\text{Exp.}\left[\frac{1}{r}\sum_{\rho=1}^{r}(y_\rho - b)^2\right] \\ &= \frac{1}{r}\int \cdots \int [(y_1 - b)^2 + \cdots + (y_r - b)^2]\, dP(y_1, \cdots y_r) \\ &= \frac{r-1}{r^2}\sum_{\rho=1}^{r}\tau_\rho^2 + \frac{1}{r}\sum_{\rho}(\beta_\rho - \beta)^2 - \frac{2}{r^2}\sum_{i<j}R_{ij}. \end{aligned} \tag{18}$$

Let us apply this result in the computation of the expectations of our test functions. It is not difficult to compute them in the general case of *different* mean values and variances. But we restrict ourselves to the consideration of certain particular cases. Take first the case where all the $m \cdot n$ mean values $\alpha_{\mu\nu}$ are equal to each other and likewise the $m \cdot n$ variances and the $\frac{1}{2}mn(mn - 1)$ covariances. Denote these magnitudes by α, σ^2 and R, respectively, we see from (18) that:

$$
\begin{aligned}
\mathcal{E}\left(\frac{s^2}{mn-1}\right) &= \mathcal{E}\left(\frac{s_a^2}{m-1}\right) = \mathcal{E}\left(\frac{\bar{s}_a^2}{n-1}\right) \\
&= \mathcal{E}\left(\frac{s_w^2}{m(n-1)}\right) = \mathcal{E}\left(\frac{\bar{s}_w^2}{n(m-1)}\right) = \mathcal{E}\left(\frac{S^2}{(m-1)(n-1)}\right) \\
&= \sigma^2 - R.
\end{aligned} \tag{19}
$$

We have thus obtained the result *that in the case of dependent populations, just described, the expectations of the six different test functions are still the same.*

Of course we may assume many other particular kinds of mutual dependence of the populations. The following assumption seems to be appropriate for problems where rows and columns play a *different* role: We consider dependence *only within each* row, that means we assume only the variables $x_{\mu 1}, x_{\mu 2}, \cdots x_{\mu n}$ as mutually dependent. The distribution (16) has then the following form:

$$V(x_{11}, \cdots x_{mn}) = V_1(x_{11}, \cdots x_{1n}) V_2(x_{21}, \cdots x_{2n}) \cdots V_m(x_{m1}, \cdots x_{mn}). \tag{20}$$

In the usual way we derive the $m \cdot n$ one dimensional distributions $V_{\mu\nu}(x)$ and the $\frac{1}{2}mn(mn - 1)$ two-dimensional distributions $V_{\mu_1\nu_1;\mu_2\nu_2}(x, y)$. If $\mu_1 \neq \mu_2$ such a two-dimensional distribution reduces to the product of the respective one-dimensional distributions. Only the $\frac{1}{2}mn(n - 1)$ bivariate distributions derived from one and the same $V_\mu(x_{\mu 1}, \cdots x_{\mu n})$ will not reduce in this way.

Denoting again by $\mathcal{E}$ the expectation with respect to $V(x_{11}, \cdots x_{mn})$ we find:

$$
\begin{aligned}
\mathcal{E}[(x_{\mu_1 i} - \alpha_{\mu_1 i})(x_{\mu_2 j} - \alpha_{\mu_2 j})] &= 0 \qquad \mu_1 \neq \mu_2 \\
&= R_{ij}^{(\mu_1)} \qquad \mu_1 = \mu_2 \text{ and } i \neq j.
\end{aligned} \tag{21}
$$

Applying now formula (18) in the computation of the expectations of s^2, s_w^2 and s_a^2 we find:

$$
\begin{aligned}
\mathcal{E}[\sum \sum (x_{\mu\nu} - a)^2] &= \frac{mn-1}{mn} \sum \sum \sigma_{\mu\nu}^2 \\
&\quad + \sum \sum (\alpha_{\mu\nu} - \alpha)^2 - \frac{2}{mn} \sum_{\mu=1}^{m} \sum_{i<j} R_{ij}^{(\mu)}, \\
\mathcal{E}[\sum \sum (x_{\mu\nu} - a_\mu)^2] &= \frac{m(n-1)}{mn} \sum \sum \sigma_{\mu\nu}^2 \\
&\quad + \sum \sum (\alpha_{\mu\nu} - \alpha_\mu)^2 - \frac{2}{n} \sum_{\mu=1}^{m} \sum_{i<j} R_{ij}^{(\mu)}, \\
\mathcal{E}[\sum \sum (a_\mu - a)^2] &= \frac{m-1}{mn} \sum \sum \sigma_{\mu\nu}^2 \\
&\quad + n \sum (\alpha_\mu - \alpha)^2 + \frac{2(m-1)}{mn} \sum_{\mu=1}^{m} \sum_{i<j} R_{ij}^{(\mu)}.
\end{aligned} \tag{22}
$$

Let us now suppose that *all the m·n distributions are equal* to each other, or, at least, that:

$$\alpha_{\mu\nu} = \alpha. \tag{23}$$

This assumption, which is characterized by (21), is, of course, different from the one which leads us to (19). We find now by means of (22), if we set

$$\sum_{\mu=1}^{m} \sum_{i<j} R_{ij}^{\mu} = \bar{R} \quad \text{and} \quad \frac{2}{mn(mn-1)} \bar{R} = R,$$

$$\mathfrak{E}\left[\frac{s_a^2}{m-1} - \frac{s^2}{mn-1}\right] = \frac{2}{mn-1} \bar{R} = \frac{mn(n-1)}{mn-1} R. \tag{24}$$

Assuming $R > 0$ (positive average correlation) we may compare this result with (11) Section 1. The term on the right side of (24) is also of the same order of magnitude as that in (11).—For negative R the term on the right side of (24) is negative and the equation may be compared with (13) Section 1. We see that for the test functions s^2/r and s_a^2/r_a *"positive, (negative) average correlation within rows" has the same effect as "Lexis (Poisson) Series" of populations.*

Consider now the test functions $\bar{s}_a^2$ and S^2. We find

$$\mathfrak{E}[\bar{s}a^2] = \mathfrak{E}[\Sigma\Sigma(\bar{a}_\nu - a)^2] = \frac{n-1}{mn} \Sigma\Sigma\sigma_{\mu\nu}^2 + m\Sigma(\bar{\alpha}_\nu - \alpha)^2 - \frac{2}{mn} \bar{R}, \tag{25}$$

and

$$\begin{aligned} \mathfrak{E}[S^2] = \mathfrak{E}[\Sigma\Sigma(x_{\mu\nu} - a_\mu - \bar{a}_\nu + a)^2] &= \frac{(m-1)(n-1)}{mn} \Sigma\Sigma\sigma_{\mu\nu}^2 \\ &+ \Sigma\Sigma(\alpha_{\mu\nu} - \alpha_\mu - \bar{\alpha}_\nu + \alpha)^2 - \frac{2(m-1)}{mn} \bar{R}. \end{aligned} \tag{25'}$$

Assuming (23) we get:

$$\mathfrak{E}\left[\frac{s_a^2}{m-1} - \frac{\bar{s}_a^2}{n-1}\right] = nR, \tag{26}$$

and if $R > 0$:

$$\mathfrak{E}\left[\frac{s_a^2}{m-1}\right] > \mathfrak{E}\left[\frac{\bar{s}_w^2}{n(m-1)}\right] > \mathfrak{E}\left[\frac{\bar{s}_a^2}{n-1}\right]. \tag{26'}$$

The first equality is analogous to (11) and (14) of Section 2 for positive or negative R respectively.[10] We also get under the assumption (23)

$$\mathfrak{E}\left[\frac{\bar{s}_a^2}{n-1}\right] = \mathfrak{E}\left[\frac{S^2}{(m-1)(n-1)}\right] = \mathfrak{E}\left[\frac{s_w^2}{m(n-1)}\right]. \tag{27}$$

---

[10] I have studied in another paper the *combination* of Lexis series and "positive correlation within rows." It turns out that the two kinds of positive effects reinforce each other. The same is true for "negative correlation" and Poisson series. See [3].

These are the same equations as (12) Section 2, and they are true for either sign of R. Hence they provide no way to decide between Lexis series and correlated populations. But computing the expectations of the magnitudes which occur in (15) Section 2 we find from (22), (25) and (25′)

$$(28)\qquad \mathcal{E}\left[\frac{s_a^2}{m-1}\right] = \sigma^2 + (n-1)R, \qquad \mathcal{E}\left[\frac{\bar{s}_w^2}{n(m-1)}\right] = \sigma^2$$
$$\mathcal{E}\left[\frac{S^2}{(m-1)(n-1)}\right] = \sigma^2 - R.$$

And hence we may say:

If the observed value of $s_a^2/(m-1)$ is greater than that of $\bar{s}_w^2/n(m-1)$ this can be explained either by the assumption of a Lexis series or a positive correlation within rows; but their equality indicate, a Poisson series; and if the first is smaller than the second we may assume negative correlation.

In the same way we may explain

$$\left[\frac{\bar{s}_w^2}{n(m-1)}\right]_{\text{observed}} > \left[\frac{S^2}{(m-1)(n-1)}\right]_{\text{observed}},$$

either by positive correlation or by Lexis series; whereas the equality indicates a Poisson series and the sign $<$ indicates negative correlation.

(*d*). *The variances of the test functions.* We have still to find the variance of our test functions. Let us compute the variance of

$$\Sigma\Sigma(x_{\mu\nu} - a_\mu - \bar{a}_\nu + a)^2$$

with respect to the $m \cdot n$ dimensional distribution $V(x_{11})V(x_{12}) \cdots V(x_{mn})$. Let us put

$$(29)\qquad x_{\mu\nu} - a_\mu - \bar{a}_\nu + a = y_{\mu\nu},$$

then we see that the average of the $y_{\mu\nu}$ equals zero

$$\bar{y} = \frac{1}{mn}\Sigma\Sigma y_{\mu\nu} = a - \frac{1}{mn} n\Sigma a_\mu - \frac{1}{mn} m\Sigma\bar{a}_\nu + a = 0.$$

and

$$S^2 = \Sigma\Sigma(x_{\mu\nu} - a_\mu - \bar{a}_\nu + a)^2 = \Sigma\Sigma(y_{\mu\nu} - \bar{y})^2.$$

Each $y_{\mu\nu}$ is a linear function of the $x_{\mu\nu}$ e.g.

$$(30)\qquad \begin{aligned} y_{11} &= x_{11}\frac{(m-1)(n-1)}{mn} \\ &\quad - \frac{m-1}{mn}\sum_{j=2}^{n} x_{1j} - \frac{n-1}{mn}\sum_{i=2}^{m} x_{i1} + \frac{1}{mn}\sum_{2}^{m}\sum_{2}^{n} x_{ij} \\ &= x_{11}\lambda_2 + \lambda_2\sum_{2}^{n} x_{1j} + \lambda_3\sum_{2}^{m} x_{i1} + \lambda_4\sum_{2}^{m}\sum_{2}^{n} x_{ij}. \end{aligned}$$

Using the same notations as in Section 1 (c) we find, because of the independence of each chance variable

$$\text{(31')}\quad \begin{aligned} \mathrm{Var}\,(y_{11}) &= \lambda_1^2\sigma^2 + \lambda_2^2(n-1)\sigma^2 + \lambda_3^2(m-1)\sigma^2 \\ &\quad + \lambda_4^2(m-1)(n-1)\sigma^2 = \frac{(m-1)(n-1)}{mn}\,\sigma^2 \end{aligned}$$

and we find the same result for each $y_{\mu\nu}$:

$$\text{(31)}\quad \sigma'^2 = \mathrm{Var}\,(y_{\mu\nu}) = \frac{(m-1)(n-1)}{mn}\,\sigma^{2\prime}$$

in agreement with the fourth line of (7) of this section. We still need M_4' the fourth moment about the mean of $y_{\mu\nu}$ which we can compute from the fourth moment of a sum. We find

$$\text{(32)}\quad M_4' = AM_4 + 6B\sigma^4,$$

and we have

$$\text{(33)}\quad \begin{aligned} A &= \lambda_1^4 + (n-1)\lambda_2^4 + (m-1)\lambda_3^4 + (m-1)(n-1)\lambda_4^4 \\ &= \frac{(m-1)(n-1)}{m^3n^3}(m^2-3m+3)(n^2-3n+3), \end{aligned}$$

and

$$\text{(34')}\quad \begin{aligned} B &= \lambda_1^2\{\lambda_2^2(n-1) + \lambda_3^2(m-1) + \lambda_4^2(m-1)(n-1)\} \\ &\quad + \lambda_2^2(n-1)\{\tfrac{1}{2}\lambda_2^2(n-2) + \lambda_3^2(m-1) + \lambda_4^2(m-1)(n-1)\} \\ &\quad + \lambda_3^2(m-1)\{\tfrac{1}{2}\lambda_3^2(m-2) + \lambda_4^2(m-1)(n-1)\} \\ &\quad + \tfrac{1}{2}\lambda_4^4(m-1)(n-1)[(m-1)(n-1)-1]. \end{aligned}$$

If we introduce the values of λ_1, λ_2, λ_3, λ_4 we find

$$\text{(34)}\quad \begin{aligned} m^4n^4B &= (m-1)^3(n-1)^3(m+n) + (m-1)^2(n-1)^2(m+n-2) \\ &\quad + \tfrac{1}{2}(m-1)(n-1)[(m-1)^3(n-2) + (n-1)^3(m-2) \\ &\quad + (mn-m-n)] \end{aligned}$$

this expression as well as that of A may be easily computed for different values of m and n.

If m and n are large, B is of order $\frac{1}{m} + \frac{1}{n}$; from (31)–(34) we see that in this case σ'^2 is approximately equal to σ^2 and M_4' to M_4.

Using now (18′) we find finally

$$\mathrm{Var}\,\{\Sigma\Sigma\,(x_{\mu\nu} - a_\mu - \bar{a}_\nu + a)^2\} = \frac{mn-1}{mn}\{(mn-1)M_4' - (mn-3)\sigma'^4\}$$

where M_4' and σ'^2 are the expressions just computed. If we compare the variances of the test functions $s_a^2/(m-1)$ and $S^2/(m-1)(n-1)$ we see that whereas the variance of the first expression is of order $1/m$ that of the second is of order $1/mn$. Hence for large values of n the *latter expression is more exact* than the former (see the analogous remark Section 1 (c)). A similar statement can be made if $\bar{s}_a^2/(n-1)$ takes the place of $s_a^2/(m-1)$.

3. Bivariate distributions. Analysis of covariance.

(*a*). *Problem.* Suppose m persons are throwing two dice, n times; we observe the respective numbers on each die in these $m \cdot n$ trials. Or we observe on m groups of n persons the color of the hair and of the eyes. Or else we state for n years the yield of wheat (in bushels) per acre and the production cost (per bushel) for m farms; etc.

We consider $m \cdot n$ pairs of numbers $x_{\mu\nu}$, $y_{\mu\nu}$. Let $V_{\mu\nu}(x, y)$[11] be the probability that $x_{\mu\nu} \leqq x$ and $y_{\mu\nu} \leqq y$; $V_{\mu\nu}(x, +\infty) = V_{\mu\nu}^{(1)}(x)$, $V_{\mu\nu}(+\infty, y) = V_{\mu\nu}^{(2)}(y)$ and introduce the following mean values and variances

$$(1) \qquad \iint x\, dV_{\mu\nu}(x, y) = \alpha_{\mu\nu}, \qquad \iint y\, dV_{\mu\nu}(x, y) = \beta_{\mu\nu},$$

$$(2) \qquad \iint (x - \alpha_{\mu\nu})^2\, dV_{\mu\nu}(x, y) = \sigma_{\mu\nu}^2, \qquad \iint (y - \beta_{\mu\nu})^2\, dV_{\mu\nu}(xy) = \tau_{\mu\nu}^2,$$

$$(3) \qquad \iint (x - \alpha_{\mu\nu})(y - \beta_{\mu\nu})\, dV_{\mu\nu}(x, y) = \gamma_{\mu\nu}$$

$$(4) \qquad \begin{aligned} &\frac{1}{n}\sum_\nu \alpha_{\mu\nu} = \alpha_\mu, \qquad \frac{1}{m}\sum_\mu \alpha_{\mu\nu} = \bar{\alpha}_\nu, \qquad \frac{1}{mn}\sum\sum \alpha_{\mu\nu} = \alpha \\ &\frac{1}{n}\sum_\nu \beta_{\mu\nu} = \beta_\mu, \qquad \frac{1}{m}\sum_\mu \beta_{\mu\nu} = \bar{\beta}_\nu, \qquad \frac{1}{mn}\sum\sum \beta_{\mu\nu} = \beta \end{aligned}$$

Let us compute the mathematical expectations of certain test functions with respect to the $2mn$-dimensional distributions

$$V_{11}(x_{11}, y_{11}) V_{12}(x_{12}, y_{12}) \cdots V_{mn}(x_{mn}, y_{mn}). \quad \text{Let}$$

$$(5) \qquad \begin{aligned} &E[F(x_{11}, y_{11}, \cdots x_{mn}, y_{mn})] \\ &\qquad = \int \cdots \int F(x_{11}, \cdots y_{mn})\, dV_{11}(x_{11}, y_{11}) \cdots dV_{mn}(x_{mn} y_{mn}) \end{aligned}$$

---

[11] In the particular case where $V_{\mu\nu}(x, y)$ has everywhere a derivative $\dfrac{\partial^2 V_{\mu\nu}}{\partial x\, \partial y}$ we can use the two dimensional density $v_{\mu\nu}(x, y) = \dfrac{\partial^2 V_{\mu\nu}}{\partial x\, \partial y}$ and the one-dimensional densities

$$v_{\mu\nu}^{(1)}(x) = \int v_{\mu\nu}(xy)\, dy; \qquad v_{\mu\nu}^{(2)}(y) = \int v_{\mu\nu}(x, y)\, dx$$

and we have

$$V_{\mu\nu}^{(1)}(x) = \int_{-\infty}^{x} v_{\mu\nu}^{(1)}(x)\, dx, \qquad V_{\mu\nu}^{(2)}(y) = \int_{-\infty}^{y} v_{\mu\nu}^{(2)}(y)\, dy.$$

We then have[12]

$$(5') \quad F[G(x_{11}, \cdots x_{mn})] = \int_{(x_{11})} \cdots \int_{(x_{mn})} G(x_{11} \cdots x_{mn})\, dV_{11}^{(1)}(x_{11}) \cdots dV_{mn}^{(1)}(x_{mn}).$$

In analogy with previous notations we introduce

$$(6) \quad \begin{aligned} a_\mu &= \frac{1}{n} \sum_\nu x_{\mu\nu}, & \bar{a}_\nu &= \frac{1}{m} \sum_\mu x_{\mu\nu}, & a &= \frac{1}{mn} \sum \sum x_{\mu\nu}, \\ b_\mu &= \frac{1}{n} \sum_\nu y_{\mu\nu}, & \bar{b}_\nu &= \frac{1}{m} \sum_\mu y_{\mu\nu}, & b &= \frac{1}{mn} \sum \sum y_{\mu\nu}, \end{aligned}$$

and

$$(7) \quad \begin{aligned} s^2 &= \Sigma\Sigma(x_{\mu\nu} - a)^2, & s_a^2 &= n\Sigma(a_\mu - a)^2, & s_w^2 &= \Sigma\Sigma(x_{\mu\nu} - a_\mu)^2 \\ S^2 &= \Sigma\Sigma(x_{\mu\nu} - a_\mu - \bar{a}_\nu + a)^2, & \bar{s}_a^2 &= m\Sigma(\bar{a}_\nu - a)^2, & \bar{s}_w^2 &= \Sigma\Sigma(x_{\mu\nu} - \bar{a}_\nu)^2 \\ t^2 &= \Sigma\Sigma(y_{\mu\nu} - b)^2, & t_a^2 &= n\Sigma(b_\mu - b)^2, & t_w^2 &= \Sigma\Sigma(y_{\mu\nu} - b_\mu)^2 \\ T^2 &= \Sigma\Sigma(y_{\mu\nu} - b_\mu - \bar{a}_\nu + b)^2, & \bar{t}_a^2 &= m\Sigma(\bar{b}_\nu - b)^2, & \bar{t}_w^2 &= \Sigma\Sigma(y_{\mu\nu} - \bar{b}_\nu)^2, \end{aligned}$$

and

$$(8) \quad \begin{aligned} c &= \Sigma\Sigma(x_{\mu\nu} - a)(y_{\mu\nu} - b), & C &= \Sigma\Sigma(x_{\mu\nu} - a_\mu - \bar{a}_\nu + a)(y_{\mu\nu} - b_\mu - \bar{b}_\nu + b) \\ c_a &= n\Sigma(a_\mu - a)(b_\mu - b) & c_w &= \Sigma\Sigma(x_{\mu\nu} - a_\mu)(y_{\mu\nu} - b_\mu) \\ \bar{c}_a &= m\Sigma(\bar{a}_\nu - a)(\bar{b}_\nu - b) & \bar{c}_w &= \Sigma\Sigma(x_{\mu\nu} - \bar{a}_\nu)(y_{\mu\nu} - \bar{b}_\nu) \end{aligned}$$

we then have

$$(9) \quad \begin{aligned} s^2 &= S^2 + s_a^2 + \bar{s}_a^2 = s_w^2 + s_a^2 = \bar{s}_w^2 + \bar{s}_a^2, \\ t^2 &= T^2 + t_a^2 + \bar{t}_a^2 = t_w^2 + t_a^2 = \bar{t}_w^2 + \bar{t}_a^2, \\ c &= C + c_a + \bar{c}_a = c_a + c_w = \bar{c}_a + \bar{c}_w, \end{aligned}$$

and corresponding relations for the ranks of these quadratic forms. We find for the expectations of these test functions, in analogy with previously investigated formulae:

$$E\left[\frac{t^2}{mn-1}\right] = \frac{1}{mn} \Sigma\Sigma\tau_{\mu\nu}^2 + \frac{1}{mn-1} \Sigma\Sigma(\beta_{\mu\nu} - \beta)^2,$$

$$E\left[\frac{t_a^2}{m-1}\right] = \frac{1}{mn} \Sigma\Sigma\tau_{\mu\nu}^2 + \frac{1}{m-1} n\Sigma(\beta_\mu - \beta)^2,$$

. . . .

and

$$E\left[\frac{c}{mn-1}\right] = \frac{1}{mn} \Sigma\Sigma\gamma_{\mu\nu} + \frac{1}{mn-1} \Sigma\Sigma(\alpha_{\mu\nu} - \alpha)(\beta_{\mu\nu} - \beta),$$

$$E\left[\frac{c_a}{m-1}\right] = \frac{1}{mn} \Sigma\Sigma\gamma_{\mu\nu} + \frac{1}{m-1} \cdot n\Sigma(\alpha_\mu - \alpha)(\beta_{\mu\nu} - \beta),$$

. . . .

[12] It may be mentioned that the problem considered in this section of mn bivariate distribution $v_{\mu\nu}(x, y)$ constitutes, of course, only a particular case of dependence (see section 2, (c)) for a $2mn$ dimensional population $v(x_{11}, y_{11}, x_{12}, y_{12}, \cdots x_{mn}, y_{mn})$.

1) *If all the* $\alpha_{\mu\nu}$ *equal each other, or all the* $\beta_{\mu\nu}$ *equal each other, we find:*

$$E_B\left[\frac{c}{mn-1}\right] = E_B\left[\frac{c_a}{m-1}\right] = E_B\left[\frac{c_w}{m(n-1)}\right]$$

$$= E_B\left[\frac{C}{(m-1)(n-1)}\right] = E_B\left[\frac{\bar{c}_a}{n-1}\right] = E_B\left[\frac{\bar{c}_w}{n(m-1)}\right] = \frac{1}{mn}\Sigma\Sigma\gamma_{\mu\nu}.$$

These formulae provide us with unbiased estimates of $\Sigma\Sigma\gamma_{\mu\nu}$.

2) *The* $\alpha_{\mu\nu}$ *are equal within each row but differ from row to row,* (Lexis) $\alpha_{\mu\nu} = \alpha_\mu \neq \alpha$; $\bar{\alpha}_\nu = \alpha$ *whereas the* $\beta_{\mu\nu}$ *may have arbitrary values, then*

$$E_L\left[\frac{\bar{c}_a}{n-1}\right] = E_L\left[\frac{c_w}{m(n-1)}\right] = E_L\left[\frac{C}{(m-1)(n-1)}\right]. \tag{13}$$

The same equalities are valid for arbitrary $\alpha_{\mu\nu}$ if the $\beta_{\mu\nu} = \beta_\mu$; $\bar{\beta}_\nu = \beta$. Our new equalities may be of some interest because inequalities analogous to those of the Lexis case cannot be proved for covariances. If the observed values of the expressions in (13) are significantly different we may conclude that neither the $\alpha_{\mu\nu}$ nor the $\beta_{\mu\nu}$ form a Lexis series. A judgment of the test (13) might be based on the investigation of its power function. But besides we have the equalities (12) and analogous equalities containing $\bar{t}_a^2$, T^2 and t_w^2.

3) *If either* $\quad \alpha_{\mu\nu} = \bar{\alpha}_\nu, \quad \bar{\alpha}_\nu \neq \alpha, \quad \alpha_\mu = \alpha,$

or $\quad \beta_{\mu\nu} = \bar{\beta}_\nu, \quad \beta_\nu \neq \beta, \quad \beta_\mu = \beta.$

We have the new equalities

$$E_P\left[\frac{c_a}{m-1}\right] = E_P\left[\frac{\bar{c}_w}{n(m-1)}\right] = E_P\left[\frac{C}{(m-1)(n-1)}\right], \tag{14}$$

and there are no inequalities analogous to the inequalities (14) *of Section* 2, *and* (13), (14) *of Section* 1.

Most of the investigations of Sections 1 and 2 can be generalized for this two dimensional problem.

BIBLIOGRAPHY

[1] R. A. Fisher, *Statistical Methods for Research Workers*, 6th ed., p. 214 ff.

[2] R. A. Fisher, "Applications of 'Student's' distributions," *Metron*, Vol. 5 (1926), pp. 90–104.

[3] H. Geiringer, "A new explanation of non-normal dispersion in the Lexis theory," *Econometrica*, Vol. 10 (1942), pp. 53–60.

[4] F. R. Helmert, *Zeits. für Math. und Physik*, Vol. 21 (1876), p. 192–218.

[5] I. O. Irwin, "Mathematical theorems involved in the analysis of variance," *Jour. Roy. Stat. Soc.*, Vol. 94 (1931), pp. 284–300.

[6] W. G. Madow, "Limiting distributions of quadratic and bilinear forms," *Annals of Math. Stat.*, Vol. 11 (1940), pp. 125–147.

[7] R. v. Mises, "Theorie des probabilites. Fondements et applications," *Annales de l'Institut Poincare*, (1931), pp. 137–190.

[8] H. L. RIETZ, "On the Lexis theory and the analysis of variance," *Bull. Am. Math. Soc.*, (1932), pp. 731 ff.

[9] A. A. TSCHUPROW, *Skandinavisk Aktuarietidskrift*, Vol. 6 (1918).

[10] A. WALD, *Lectures on the Analysis of Variance and Covariance*, *Columbia University*, 1941.

[11] MILTON FRIEDMAN, "The use of ranks to avoid the assumption of normality," *Jour. Amer. Stat. Assn.*, Vol. 32 (1937), pp. 675–701.