KONINKLIJKE NEDERLANDSCHE AKADEMIE VAN WETENSCHAPPEN

Construction of a confidence region for a line

BY

J. HEMELRIJK

Reprinted from Proceedings Vol. LII, No. 9, 1949

Reprinted from Indagationes Mathematicae, Vol. XI, Fasc. 5, 1949

1949
NORTH-HOLLAND PUBLISHING COMPANY
(N.V. Noord-Hollandsche Uitgevers Mij.)
AMSTERDAM

J. HEMELRIJK: *Construction of a confidence region for a line.*

(Communicated at the meeting of September 24, 1949.)

1 *Introduction:*

Let Γ be a *probability set* („Wahrscheinlichkeitsfeld" according to A. KOLMOGOROFF), i.e. a set Γ of elements λ, upon which an absolutely additive set function F is given (defined for all subsets Λ of Γ belonging to a given closed family H of subsets which contains Γ), with the properties

$$F(\Lambda) \geqq 0 \quad \text{for every} \quad \Lambda \in H$$
$$F(\Gamma) = 1.$$

Then a *random variable* $\underline{x}$ [1]) can be considered as a function, defined for every $\lambda \in \Gamma$ and taking there the value $x(\lambda)$. If Λ is the subset of Γ, where $x(\lambda)$ takes a certain set X of values, the probability that $\underline{x} \in X$ (denoted by $P[\underline{x} \in X]$) is

$$P[\underline{x} \in X] = F(\Lambda). \quad \ldots \ldots \quad (1)$$

A *random element* $\underline{\varphi}$ of some set K (e.g. a random point or a random vector) can analogously be defined by adjoining an element φ of K to every element λ of a probability set Γ (notation: $\varphi(\lambda)$); and a *random system* $\underline{\Phi}$ *of elements* φ by adjoining a subset Φ of K to every element λ of a probability set Γ (notation: $\Phi(\lambda)$).

If $\underline{\Phi}$ is a random system of elements φ, and if φ_0 is one such element; if furthermore the random variable $\underline{u}(\varphi_0)$ is defined by the relations

$$u(\varphi_0; \lambda) = 1 \quad \text{if} \quad \varphi_0 \in \Phi(\lambda)$$
$$u(\varphi_0; \lambda) = 0 \quad \text{if not,}$$

then $\underline{\Phi}$ is called a *confidence region* for φ_0 with *confidence level* $p = P[\underline{u}(\varphi_0) = 0]$ (or: *confidence coefficient* $1 - p = P[\underline{u}(\varphi_0) = 1]$).

2. *The problem.*

This may be formulated as follows:
Given: 1. a random set of n points $\underline{P}_i$ $(i = 1, \ldots, n)$ in a plane V satisfying the relations

$$\underline{P}_i = Q_i + \underline{\mathbf{w}}_i \qquad (i = 1, \ldots, n) \quad \ldots \ldots \quad (2)$$

[1]) The random character of a variable, or, in general, of an element of some set, will be indicated by underlining the symbol, which denotes the variable or element respectively.

where $\underline{w}_i$ is a random system of vectors in V [2]) and $Q_1, \ldots, Q_n$ are fixed points in V, situated on a straight line L, given by the equation

$$L \equiv \alpha\xi + \beta\eta + \gamma = 0 \quad \ldots \ldots \ldots \quad (3)$$

where ξ and η are Cartesian coordinates in V;

2. some conditions, which will be specified later, about the probability distribution of the random system of errors $\underline{w}_i$;

3. a real number p with $0 < p < 1$;

To find: a confidence region $\underline{R}$ for L, consisting of lines in V and depending on $\underline{P}_1, \ldots, \underline{P}_n$ only, with confidence level $\leqq p$.

In particular we shall give a construction depending on $\underline{P}_1, \ldots, \underline{P}_n$ only, of confidence regions for:

I. The direction δ_0 of L $(\delta_0 = -\alpha/\beta)$

II. The intercept $\tau_0 = -\gamma/\beta$ of L, a) under the condition $\delta_0 = \delta$ and b) unconditionally

III. δ_0 and τ_0 jointly.

The construction of a joint confidence region for δ_0 and τ_0 is equivalent with the construction of $\underline{R}$. The constructions will be given in separate sections, the conditions concerning the probability distribution of the random set of errors $\underline{w}_i$ being mentioned at the beginning of each section.

The probability set of the random system of errors $\underline{w}_i$ $(i = 1, \ldots, n)$, for which we may take a $2n$-dimensional Cartesian space, will be called Γ. Each element $\lambda \in \Gamma$ then corresponds with a specified system of errors $w_i(\lambda)$ and, Q_i being fixed, with a specified system of points $P_i(\lambda)$ $(i = 1, \ldots, n)$. Therefore Γ may be taken as the probability set of the random system of points $\underline{P}_i$ also.

3. *Remarks.*

3.1. The problem under consideration may arise in many fields of science, e.g. in physics, chemics and economics. If ξ and η are two variables, known (or supposed) to be linearly connected according to equation (3) with unknown coefficients α, β and γ; if furthermore the measurements of both ξ and η are subject to error; then the determination of a joint confidence-region for $-\alpha/\beta$ and $-\gamma/\beta$ by means of n observed points P_i with coordinates (x_i, y_i) $(i = 1, \ldots, n)$ is identical with our problem. To every observed point P_i a "true", but unknown, point Q_i is then supposed to correspond, according to the equations (2), where w_i represents the error of the i^{th} observation. To these errors w_i corresponds an element $\lambda \in \Gamma$ (where Γ is the probability set of the errors, cf. 2.) and to this element λ corresponds a set $R(\lambda)$ of lines in V, which can be con-

[2]) The vectors $\underline{w}_i$ will be called the errors.

structed by means of the points $P_1, \ldots, P_n$. $R(\lambda)$ may be regarded as an "observation" of the random confidence region $\underline{R}$ for L, corresponding to the observed points $P_1, \ldots, P_n$. The property, that the confidence coefficient is $1-p$ is then usually expressed by saying, that the probability, that L is an element of $R(\lambda)$ is equal to $1-p$.

3.2. A solution for a special case has been given by A. WALD in 1940 [3]). The conditions he imposes on the random set of errors are, however, rather more stringent (e.g. normality) than the conditions, which will be used here. In the same paper he derives consistent estimates of the coefficients $-\alpha/\beta$ and $-\gamma/\beta$ under less stringent conditions. We shall show that fairly general conditions are sufficient for the construction of confidence regions of these coefficients.

The methods used are different from those used by WALD, and of an elementary nature. They are related to those generally employed for the parameter-free construction of confidence intervals. The smallest number of points, which is needed for the construction of the confidence-regions mentioned above, with a reasonably large confidence coefficient (about 0,95) will prove to be seven.

Another partial solution of our problem, together with the solution of some other problems, has recently been found by H. THEIL. In particular he gives another confidence region for $-\alpha/\beta$ under conditions of the same nature as those imposed here, in a publication shortly to appear.

4. *Confidence region* $\underline{D}$ *for the direction* δ_0 *of* L.

4.1. *Condition* I:

a. The n random errors $\underline{\mathbf{w}}_i$ $(i = 1, \ldots, n)$ are independently distributed with twodimensional probability distributions, which are the same for every i.

b. If $\underline{u}_i$ and $\underline{v}_i$ are the components of $\underline{\mathbf{w}}_i$ in the direction of the ξ- and η-axes of V, then the probability, that the random point with coordinates $\underline{u}_i$ and $\underline{v}_i$ lies on a fixed straight line N in V is equal to zero for every N in V (and for every i) [4]).

Remark: strictly speaking it is sufficient if condition I *b* is fulfilled for all lines parallel to L only. In general however, L being unknown, this amounts to the same as I *b*.

4.2. *Notation:* We shall call the strip (including its boundaries) of the plane V, bounded by two parallel straight lines through P_r and P_s $(r \neq s)$ and having the direction δ the $(r, s; \delta)$*-strip*. When $\underline{P}_r$ and $\underline{P}_s$ are random points, this strip is a random strip with fixed direction δ.

[3]) The fitting of straight lines if both variables are subject to error, Ann. Math. Stat. 11 p. 284—300 (1940). This paper contains a summary of earlier results.

[4]) Interpreting a probability distribution as the distribution of a unit mass over the probability set, this means, that no straight line bears a positive mass.

A direction δ will be called $(r, s; m)$-*rejectable* with respect to the specified system of points P_i $(i = 1, \ldots, n)$, if the $(r, s; \delta)$-strip corresponding to $P_1, \ldots, P_n$, contains at least $n-m$ of the points $P_1, \ldots, P_n$ and $(r, s; m)$-*acceptable* (where "acceptable" is short for "non-rejectable") if this is not the case.

The absolutely additive set function F on the probability set Γ (cf. 2) then determines the probability, that a fixed direction δ will be $(r, s; m)$-rejectable for given r and s: if $\Lambda \subset \Gamma$ is the subset of those λ, for which δ is $(r, s; m)$-rejectable, then $F(\Lambda)$ is this probability.

4. 3. *Theorem* I:

If $\underline{P}_i = Q_i + \underline{w}_i$ $(i = 1, \ldots, n)$ *are n random points in a plane V, where* $Q_i, \ldots, Q_n$ *lie on a straight line L in V and* $\underline{w}_1, \ldots, \underline{w}_n$ *fulfill condition I, then for any fixed r and s (with* $r \neq s$; $1 \leqq r \leqq n$; $1 \leqq s \leqq n$) *and for any natural number m* $(0 \leqq m \leqq n-2)$, *the set* $\underline{D}$ *of* $(r, s; m)$*-acceptable directions is a confidence region for the direction* δ_0 *of L, with confidence level*

$$p_1 = \frac{(m+1)(m+2)}{n(n-1)} \qquad (0 \leqq m \leqq n-2). \quad . \quad . \quad . \quad . \quad (4)$$

Proof: To prove the theorem, we only have to show that the probability, that δ_0 is $(r, s; m)$-rejectable is equal to p_1.

Now δ_0 is $(r, s; m)$-rejectable, if and only if the $(r, s; \delta_0)$-strip contains at least $n-m$ of the points $\underline{P}_1, \ldots, \underline{P}_n$. Denoting the distance from $\underline{P}_i$ to L, measured in an arbitrary fixed direction different from δ_0, by $\underline{z}_i$, this means, that at least $n-m$ of the quantities $z_1, \ldots, \underline{z}_n$ lie in the closed interval $(\underline{z}_r, \underline{z}_s)$. According to condition I the $\underline{z}_i$ $(i = 1, \ldots, n)$ are distributed independently, according to a probability distribution, which is the same for every i and which is continuous because of condition I *b*. Therefore the probability is equal to one, that all $\underline{z}_i$ are different and, arranging them according to decreasing magnitude, $\underline{z}_r$ has, for every j, probability $1/n$ to be the jth one from the top. If δ_0 is $(r, s; m)$-rejectable, $\underline{z}_r$ must have one of the $m+1$ largest or one of the $m+1$ smallest values and if it takes the jth value (with $j \leqq m+1$) from the top (or from the bottom respectively), then $\underline{z}_s$ must take one of the $m+2-j$ smallest (or largest) values respectively. The probability, that δ_0 is $(r, s; m)$-rejectable is therefore equal to

$$2 \sum_{j=1}^{m+1} \frac{1}{n} \cdot \frac{m+2-j}{n-1} = \frac{(m+1)(m+2)}{n(n-1)}$$

which is equal to p_1.

Remark: $\underline{D}$ consists of a finite number of angles [5]) corresponding with a finite number of intervals for $-\alpha/\beta$. It reduces to one angle if there is

[5]) Where "angle" stands for "pair of vertically opposite angles".

a $(r, s; \delta)$-strip, which contains all points $P_1, \ldots, P_n$. This condition is sufficient, but not necessary.

4. 4. *On the choice of the numbers r and s.*

Theorem I has been proved without imposing any restrictions on the choice of $\underline{P}_r$ and $\underline{P}_s$ out of $\underline{P}_1, \ldots, \underline{P}_n$. It must, however, be pointed out, that this choice must be independent of the vectors $\underline{\mathbf{w}}_i$, because otherwise, the $\underline{z}_i$ $(i = 1, \ldots, n)$ do not necessarily have the same probability distribution any more.

Bearing this restriction in mind, we now consider the question, which choice would, on the average, be preferable. It is clear that, unless the points $\underline{P}_i$ lie on a straight line (in which case every direction is $(r, s; m)$-rejectable; this case, however, has probability zero), the direction $\underline{d}_{rs}$ of the line $\underline{P}_r\underline{P}_s$ is always $(r, s; m)$-acceptable. Clearly, the method will gain in power, if we do not choose r and s arbitrarily, but if we choose them so, that the probability of a large deviation of $\underline{d}_{rs}$ from the direction δ_0 of L, is minimised. A choice of r and s, which attains this end for every deviation, will therefore be considered preferable.

Supposing the indices of the points Q_i $(i = 1, \ldots, n)$ to be chosen such, that $Q_1 \neq Q_n$ and that Q_i for $i = 2, \ldots, n-1$ lies in the open interval (Q_1, Q_n), it is easy to see, that the choice $r = 1$ and $s = n$ (or $s = 1$, $r = n$) is preferable in the abovementioned sense to all other choices.

To prove this, we consider, for every r and s with $r \neq s$, the two-dimensional probability set N_{rs} of the random system of the two vectors $\underline{\mathbf{w}}_r$ and $\underline{\mathbf{w}}_s$. N_{rs}, as well as the absolutely additive set function on N_{rs} representing the joint probability distribution of $\underline{\mathbf{w}}_r$ and $\underline{\mathbf{w}}_s$, are the same for every r and s $(r \neq s)$. Every element μ of N_{rs} corresponds with a deviation $\triangle_{rs}(\mu)$ of $d_{rs}(\mu)$ from δ_0. This deviation $\triangle_{rs}(\mu)$, however, is, for every element μ of N_{rs}, smallest if $r = 1$ and $s = n$ (or $r = n$ and $s = 1$), which follows easily from the fact, that Q_1 and Q_n are the extreme point of $Q_1, \ldots, Q_n$. This proves our contention.

A second reason, for preferring the choice $r = 1$ and $s = n$ is, that, according to the remark of the preceding section, the probability that $\underline{D}$ consists of a single angle, is then as large as possible.

In general it will not be possible to select from a specified system of points P_i $(i = 1, \ldots, n)$ the points P_1 and P_n corresponding to Q_1 and Q_n, without making any further assumptions about the errors, because the points Q_i are unknown. It may occur, that the points P_1 and P_n can be selected on non-statistical considerations, for instance if it is known, that the points $Q_1, \ldots, Q_n$ form a monotonous sequence, being observed in the same order (e.g. if ξ denotes the time when the observation takes place). Another situation, which may arise is, that we have a criterion C at our disposal, which (under some further assumptions for the errors) indicates unambiguously among every specified system of points $P_1, \ldots, P_n$ (except perhaps with zero probability) two points P_r and P_s with $r \neq s$ and with the property, that

$$P[(r=1 \text{ and } s=n) \text{ or } (r=n \text{ and } s=1)] \geqq 1-q. \quad . \quad . \quad (5)$$

C may, for instance, consist of taking the point P_i with smallest abscissa as P_r and the one with largest abscissa as P_s. We shall not occupy ourselves with a discussion of the different possibilities for C and the computation of the corresponding q, this being quite a subject in itself [6]). We only point out, that, if (5) is valid, theorem I remains correct with confidence level $p_1^* \leqq p_1 + q - p_1 q$, if for $\underline{P}_r$ and $\underline{P}_s$ we take the points indicated by C, instead of keeping r and s constant. This may be seen as follows: denoting by A the event, that C has indicated the right pair of points, we have

$$P[\delta_0 \in D | A] = 1 - p_1$$

i.e. the conditional probability, under the condition A, that $\delta_0 \in D$, is equal to $1 - p_1$. Thus:

$$P[\delta_0 \in D] \geqq P[A \text{ and } \delta_0 \in D] = P[A] \cdot P[\delta_0 \in D | A] \geqq (1-q)(1-p_1)$$
$$p_1^* = 1 - P[\delta_0 \in D] \leqq 1 - (1-q)(1-p_1) = p_1 + q - p_1 q.$$

5. *Conditional confidence interval* $\underline{T}$ *for* $\tau_0 = -\gamma/\beta$ *under the condition* $\delta_0 = \delta$.

5. 1. *Condition* II: The random vectors $\underline{w}_i$ $(i = 1, \ldots, n)$ are distributed independently; for every i the distribution of $\underline{w}_i$ is such, that the random point $\underline{P}_i$ has equal probability to lie on either side of L and probability zero to lie on L.

Remark: The distribution of $\underline{w}_i$ may now depend on i. Condition II is satisfied if e.g. the distribution of $\underline{w}_i$ (for every i) is symmetrical with respect to the origin.

5. 2. *Notation:* We shall call τ a (δ, k)-*rejectable* value of the intercept, under the condition $\delta_0 = \delta$, with respect to a specified system of points $P_1, \ldots, P_n$, if at most k of the points P_i $(i = 1, \ldots, n)$ are situated on one side of the line L' through the point $(0, \tau)$ with direction δ.

If on both sides of L' lie more than k points, τ will be called (δ, k)-*acceptable* under the condition $\delta_0 = \delta$ (where again "acceptable" is short for "non-rejectable"). The condition $\delta_0 = \delta$ will not always be mentioned explicitly.

The absolutely additive set function F on Γ then determines the pro-

---

[6]) It is clear that especially if the length of $\underline{w}_i$ has a finite range, q will be equal to 0, if the distance of the points with smallest and largest abscissae (or ordinates) have a distance larger than *four* times this range to all other points.

bability, that a fixed value τ will be (δ, k)-rejectable: if $\Lambda \subset \Gamma$ is the subset of those λ, for which τ is (δ, k)-rejectable, then $F(\Lambda)$ is this probability.

5.3. *Theorem* II: *If* $\underline{P}_i = Q_i + \underline{\mathbf{w}}_i$ $(i = 1, \ldots, n)$ *are n random points in a plane V, where* $Q_1, \ldots, Q_n$ *lie on a straight line L in V and* $\underline{\mathbf{w}}_1, \ldots, \underline{\mathbf{w}}_n$ *fulfill condition II, then the set* $\underline{T}$ *of* (δ, k)*-acceptable values* τ *(where k is an integer* $< \frac{n-3}{2}$*) of the intercept is a conditional confidence interval for the intercept* τ_0 *under the condition* $\delta_0 = \delta$*, with confidence level*

$$p_2 = 2^{-n+1} \sum_{i=0}^{k} \binom{n}{i} \qquad \leqq k < \tfrac{n-3}{2}) \; . \; . \; . \; . \; . \; . \; . \; . \quad (6)$$

Proof: From the definition of a (δ, k)-rejectable value τ it is clear, that the set $\underline{T}$ is a random interval. Remains to calculate the confidence level. For every point $\underline{P}_i$ the probability to lie at either side of L is equal to $\frac{1}{2}$. Therefore the probability, that at one of the sides of L lie k or less points $\underline{P}_i$, is equal to

$$2^{-n+1} \sum_{i=0}^{k} \binom{n}{i} .$$

6. *Confidence region* $\underline{R}$ *for L.*

6.1. *Condition* III: conditions I and II are both satisfied; i.e.: *a.* The errors $\underline{\mathbf{w}}_i$ are independently distributed, with two dimensional probability distributions, which are the same for every i.

b. The probability, that $(\underline{u}_i, \underline{v}_i)$, where $\underline{u}_i$ and $\underline{v}_i$ are the components of $\underline{\mathbf{w}}_i$, lies on a fixed straight line parallel to L, is equal to zero, for every such line.

c. The probability, that $\underline{P}_i$ lies above L is equal to $\frac{1}{2}$. The probability, that $\underline{P}_i$ lies on L is equal to zero.

6.2. *Theorem* III: *If* $\underline{P}_i = Q_i + \underline{\mathbf{w}}_i$ $(i = 1, \ldots, n)$ *are n random points in a plane V, where* $Q_1, \ldots, Q_n$ *lie on a straight line L in V and* $\underline{\mathbf{w}}_1, \ldots, \underline{\mathbf{w}}_n$ *fulfill condition III; if r and s are two different integers taken from* $1, \ldots, n$*; if m is an integer with* $0 \leqq m \leqq n-2$ *and if k is an integer with* $0 \leqq k < \frac{n-3}{2}$ *; then the set* $\underline{R}$ *consisting of those lines in V of which both the direction* δ *is* $(r, s; m)$*-acceptable and the intercept* τ *is* (δ, k)*-acceptable, is a confidence region for L with confidence level*

$$p = p_1 + p_2 - p_1 p_2 \; . \; . \; . \; . \; . \; . \; . \; . \; . \quad (7)$$

where

$$p_1 = \frac{(m+1)(m+2)}{n(n-1)} \quad \text{and} \quad p_2 = 2^{-n+1} \sum_{i=0}^{k} \binom{n}{i}.$$

Proof: The proof consists again of showing, that the set $\underline{R}$ has probability p not to contain L. According to theorem I and II, the probabilities, that δ_0 is $(r, s; m)$-rejectable and that τ_0 is (δ_0, k)-rejectable are respectively p_1 and p_2. Now the $(r, s; m)$-rejectability of δ_0 depends on the place which $\underline{z}_r$ and $\underline{z}_s$ (cf. the proof of theorem I) take in the sequence of $\underline{z}_1, \ldots, \underline{z}_n$ when arranged according to decreasing magnitude. The (δ_0, k)-rejectability of τ_0, however, is invariant against permutations of the points $\underline{P}_i$; hence the $(r, s; m)$-rejectability of δ_0 and the (δ_0, k)-rejectability of τ_0 are independent. From this (7) follows.

6.3. *The actual construction of* $\underline{R}$.

In diagram 1 an example is given of the form which the set $\underline{R}$ can take in a specified case (i.e. for one element λ of Γ). P_r and P_s have been supposed to be the points with smallest and largest abscissa (cf. 4.4; if e.g. the error in the ξ-direction is sufficiently small in comparison with the differences of the abscissae of these points and the other points, this procedure is justified).

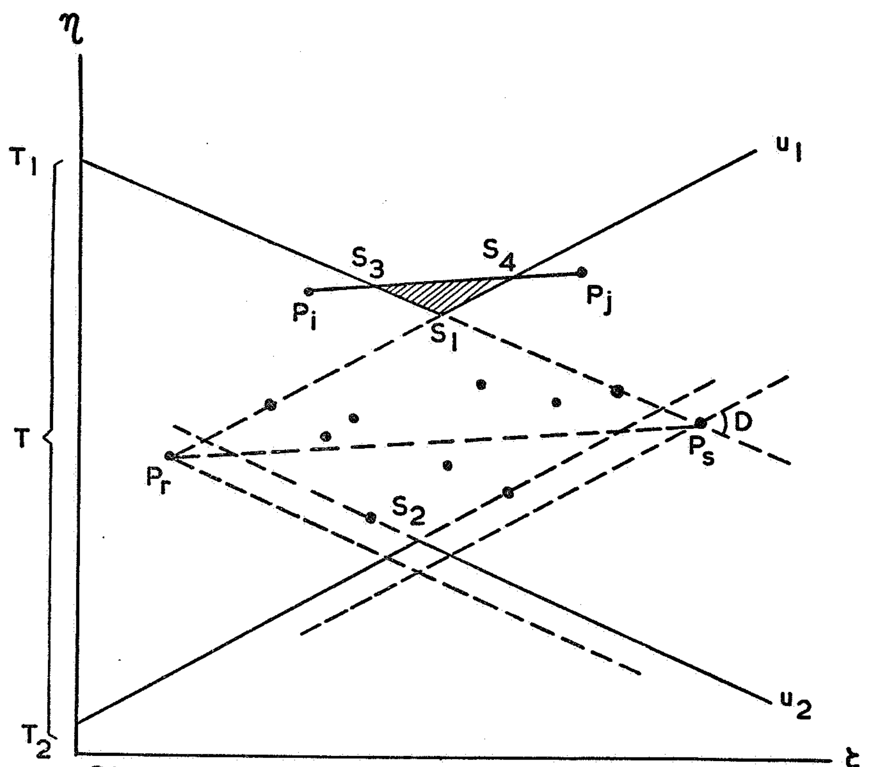

Diagram 1, n = 13, m = 1, k = 1.

First D is constructed by letting two parallel lines revolve around P_r and P_s respectively and registering the $(r, s; m)$-acceptable directions. Then the parallel lines through P_r and P_s in both extreme acceptable directions are pushed together (or eventually pulled apart) untill a position is reached, where they indicate the extreme lines of their own direction δ which are (δ, k)-acceptable. This gives the "diabolo" $T_1 S_1 U_1$; $T_2 S_2 U_2$.

Next all points P_i lying outside one of the strips bounded by $T_1 S_1$ and $S_2 U_2$ or $T_2 S_2$ and $S_1 U_1$ respectively are connected by straight lines. From these those lines are selected, which have an $(r, s; m)$-acceptable direction δ, and which have an intercept τ which is (δ, k)-acceptable or on the verge of (δ, k)-acceptability (like $P_i P_j$ in the diagram). The portions (like $S_3 S_1 S_4$) which these lines cut of from $T_1 S_1 U_1$ and $T_2 S_2 U_2$ are joined to the diabolo.

The resulting region of the plane V then contains all lines of R; a line however, lying in this region, does not necessarily belong to R because, although its direction δ is acceptable, its intercept may be (δ, k)-rejectable.

In the diagram we have $n = 13$, $m = k = 1$; hence $p = 0.043$.

The construction can easily be carried out graphically by taking e.g. a very large η-scale, so that the ordinates of the points P_i have a large variation.

7. *Miscellanous remarks.*

7.1. *Unconditional confidence interval for τ_0.*

The set of those points of the η-axis, which lie on a line of $\underline{R}$, is a confidence interval for τ_0 with confidence level p, without condition about the direction of L. In diagram 1 this interval is (T_1, T_2).

7.2. *Conditional confidence region for δ_0 under the condition $\tau_0 = \tau$.*

This confidence region consists of the direction of those lines through the point $(0, \tau)$, for which:

1. The direction δ is $(r, s; m)$-acceptable.
2. τ is (δ, k)-acceptable.

The confidence level then is p.

The set of directions δ of those lines through the point $(0, \tau)$, for which τ is (δ, k)-acceptable, is another conditional confidence region for δ_0, containing the first one, with confidence level p_2.

7.3. *Testing of hypotheses.*

From the foregoing sections simple tests can be derived for the hypotheses, *a*) that L is a given line L', and *b*) that L contains a given point Q_0.

The test of the hypothesis $L' = L$ consists of drawing two lines L_1' and L_2' parallel to L' through P_r and P_s and counting the number of points P_i outside the strip bounded by L_1' and L_2'. Calling this number m' and calling the numbers of points P_i lying on the two sides of L', k' and k'' respectively, the hypothesis $L' = L$ is rejected if either $m' \leqq m$ or $\text{Min}\,(k', k'') \leqq k$ (i.e. if L' does not belong to R). The level of significance of this test is $p = p_1 + p_2 - p_1 p_2$.

In an analogous way the other hypothesis mentioned may be tested without carrying out the complete construction of R.

7.4. *Table of p_1 and p_2.*

n ↓ \ m, k →	0	1	2	3	4	5
6	0.067 0.032					
7	0.048 0.016					
8	0.036 0.008	0.108 0.071				
9	0.028 0.004	0.084 0.040				
10	0.022 0.002	0.067 0.022				
11	0.018 0.001	0.055 0.012	0.110 0.066			
12	0.015 0.0005	0.046 0.007	0.091 0.039			
13	0.013 0.0003	0.039 0.004	0.077 0.023	 0.093		
14	0.011 0.0002	0.033 0.002	0.066 0.013	 0.058		
15	0.010 0.00006	0.029 0.001	0.058 0.008	0.096 0.036		
16	0.009 0.00004	0.025 0.0006	0.050 0.005	0.084 0.022	 0.077	
17	0.008 0.00002	0.023 0.0003	0.045 0.003	0.074 0.013	 0.050	
18	0.007 0.000008	0.020 0.0002	0.040 0.002	0.066 0.008	0.099 0.031	 0.097
19	0.006 0.000004	0.018 0.00008	0.036 0.0008	0.059 0.005	0.088 0.020	 0.064
20	0.006 0.000002	0.016 0.00005	0.032 0.0005	0.053 0.003	0.079 0.012	 0.042

$$p_1 = \frac{(m+1)(m+1)}{n(n-1)} \qquad p_2 = 2^{-n+1} \sum_{i=0}^{k} \binom{n}{i}.$$

In every compartment the number at the top represents p_1 and the number at the bottom p_2; p_1 and p_2 need not be taken from the same partition in a row.

Values of p_1 have been included up to about 0.10 and of p_2 to such a level, that with the same n there is a p_1 which makes $p_1 + p_2$ not larger than about 0.10; the reason for including these rather high values is, that in special cases regions may be indicated corresponding with a confidence level of $\frac{1}{2}(p_1 + p_2) - \frac{1}{4} p_1 p_2$. In the diagram of 6. 3 e.g. the part of the η-axis above T_1 contains τ_0 with this probability only, if the error in the ξ-direction is so small, that the abscissa of P_r must necessarily be smaller than the abscissa of P_s for every $\lambda \in \Gamma$, and that, at the same time, no point P_i has a negative abscissa for any $\lambda \in \Gamma$. The same property then holds for the part of the η-axis below T_2. We omit the proof of this contention; it runs along the same lines as the proofs of the other theorems, applying one-sided criteria for rejectability instead of the two-sided criteria, which have been used there.

I want to thank Prof. Dr D. VAN DANTZIG, whose suggestions helped me to give the paper its final form.

Publication of the Statistical Department of the "Mathematisch Centrum", Amsterdam.