ON FAMILIES OF ADMISSIBLE TESTS

By E. L. Lehmann

University of California, Berkeley

1. Summary. For each hypothesis H of a certain class of simple hypotheses, a family F of tests is determined such that

(a) given any test w of H there exists a test w' belonging to F which has power uniformly greater than or equal to that of w.

(b) no member of F has power uniformly greater than or equal to that of any other member of F.

The effect on F of various assumptions about the set of alternatives are considered. As an application an optimum property of the known type A_1 tests is proved, and a result is obtained concerning the most stringent tests of the hypotheses considered.

2. Introduction. In the theory of testing simple hypotheses, if a uniformly most powerful test exists, it is the most desirable test to use. If, as is generally the case, such a test does not exist, the choice between tests none of which is "altogether better" than all the others, has to be based on information not contained in the general formulation of the testing problem. If no such additional information is available, the choice must of necessity be somewhat arbitrary.

Now although a single uniformly most powerful test exists only in exceptional cases, there will always exist a family F of tests such that

(a) given any test w of the hypotheses H under consideration and of prescribed level of significance, there exists a test w' belonging to F which has power uniformly greater than or equal to that of w.

(b) no member of F has power uniformly greater than or equal to that of any other member of F.

The family F is essentially unique. Arbitrariness occurs only since a test region is not uniquely determined by its power function. But since two tests with the same power function are equivalent for testing purposes, it is from the present point of view immaterial which one is included in F.

With the same restriction F is essentially the family of admissible tests, a test w being admissible if there is no test of the same level of significance which has power uniformly greater than or equal to but not identically equal to that of w. This definition differs only trivially from the one given by Wald [1, p. 15] who defines a test w to be non-admissible if there exists a w' with power everywhere greater than that of w (except at the hypothetical point).

F naturally depends on the class of alternatives considered. A restriction in the class of alternatives may (although it will not necessarily) diminish F. The family F may also be decreased by other additional information: For instance a probability distribution may be assumed for the set of alternatives, and some properties of this distribution may be presupposed.

The determination of the family F, (and a description of the power functions of the tests in F) might be considered a solution of the testing problem. The solution is not unique and hence does not provide a basis for action. This reflects the fact that additional information is needed to make possible the unique choice of a best test. On the basis of the available information, F represents the furthest reduction of the problem that seems possible. On the one hand, if the choice of test is to be made from the point of view of power, the only contestants for "best test" are the members of F. On the other hand, the available information does not give preference to any one member of F over any other unless additional principles (such as unbiasedness for instance) are introduced.

It is the purpose of the present paper to illustrate the above notions by determining F for a very simple case.

3. Determination of the family F.

Let the random variable

$$E = (X_1, X_2, \cdots; X_n)$$

have a probability density function

(1)
$$p_\theta$$

depending on parameter θ. Concerning (1) we shall make the assumptions under which Neyman [2, 3] has shown the existence of the type A_1 test of the hypothesis

(2)
$$H: \theta = \theta_0.$$

ASSUMPTIONS:

(a) *Conditions of regularity:*
The integral

(3)
$$\int_w p_\theta(e)\, de \qquad \begin{aligned} e &= (x_1, \cdots, x_n) \\ de &= dx_1 \cdots dx_n \end{aligned}$$

extended over any region w in the sample space, admits of two successive derivatives with respect to θ under the integral sign, i.e.

(4)
$$\frac{d^k}{d\theta^k} \int_w p_\theta(e)\, de = \int_w \frac{\partial^k}{\partial\theta^k} p_\theta(e)\, de \quad \text{for} \quad k = 1, 2.$$

(b) *A differential equation:*
If

(5)
$$\varphi_\theta(e) = \frac{\partial}{\partial\theta} \log p_\theta(e)$$

$$\varphi_\theta'(e) = \frac{\partial}{\partial\theta} \varphi_\theta(e),$$

φ_{θ_0} is not identically zero, and there exist functions of θ (but independent of e), A and B, such that

(6)
$$\varphi_\theta' = A + B\varphi_\theta .$$

Under these assumptions Neyman has shown

A. that the probability density function p_θ is of the form

(7)
$$p_\theta(e) = \exp \{P(\theta) + T(e) \cdot Q(\theta) + R(e)\}$$

where Q is a monotone function with $\frac{d}{d\theta} Q(\theta) \mid_{\theta=\theta_0} \neq 0$ (without loss of generality we shall assume Q monotonely increasing) and

B. that the type A_1 test of the hypothesis H exists, and is given by

(8)
$$T(e) \leq c_1 , \qquad T(e) \geq c_2$$

for suitable choice of c_1 and c_2 .

In what follows we shall assume that the permissible first kind error in testing H is fixed throughout and has the value ϵ. By a test w of H we shall always mean a test of level of significance ϵ, i.e. satisfying

(9)
$$\int_w p_{\theta_0}(e) \, de = \epsilon.$$

Let us consider the family of tests

(10)
$$w(k)\colon T(e) \leq k, T(e) \geq f(k); k < f(k)$$

where $f(k)$ is determined by (9). It easily follows from (9) that k can take on all values from $-\infty$ to k_0 , say, where k_0 is such that

(11)
$$f(k_0) = +\infty .$$

For the family F of tests $\{w(k)\}$, $-\infty \leq k \leq k_0$ we now state

THEOREM 1. *All members of* F *are admissible, and if* w *is any admissible test not in* F, *there exists a member of* F *which has power identical with that of* w.

We first prove the

LEMMA. *Let* β_w *denote the powerfunction of a test* w. *Then if* $k_1 < k_2$

(12)
$$\beta_{w(k_1)}(\theta) < \beta_{w(k_2)}(\theta) \qquad\qquad \text{if } \theta < \theta_0$$
$$\beta_{w(k_1)}(\theta) > \beta_{w(k_2)}(\theta) \qquad\qquad \text{if } \theta > \theta_0 .$$

PROOF: Let $\bar{w}$ denote the complement of a region w. Consider the intervals

(13)
$$I = w(k_1) \cdot \overline{w(k_2)}$$
$$J = \overline{w(k_1)} \cdot w(k_2).$$

I lies entirely to the right of J. Let $\theta > \theta_0$. Then

(14)
$$\frac{p_\theta(e)}{p_{\theta_0}(e)} = C(\theta) \exp \{T(e)[Q(\theta) - Q(\theta_0)]\}$$

is a strictly increasing function of T since Q is increasing. Therefore there exists a constant C such that

(15)
$$\frac{p_\theta(e)}{p_{\theta_0}(e)} \le C \quad \text{if} \quad T(e) \text{ is in } J$$

$$C \le \frac{p_\theta(e)}{p_{\theta_0}(e)} \quad \text{if} \quad T(e) \text{ is in } I.$$

Since

(16)
$$\int_{w(k_1)} p_{\theta_0}(e) \, de = \int_{w(k_2)} p_{\theta_0}(e) \, de$$

we have

(17)
$$\int_{T(e) \epsilon I} p_{\theta_0}(e) \, de = \int_{T(e) \epsilon J} p_{\theta_0}(e) \, de$$

and therefore

(18)
$$\int_J p_\theta(e) \, de < C \cdot \int_J p_{\theta_0}(e) \, de = C \cdot \int_I p_{\theta_0}(e) \, de < \int_I p_\theta(e) \, de$$

from which it follows that

(19)
$$\int_{w(k_2)} p_\theta(e) \, de < \int_{w(k_1)} p_\theta(e) \, de$$

which is the desired result.

PROOF OF THEOREM 1. The proof consists of several parts.

I. Let m be any real number, and assume that there exists a value of k such that

(20)
$$\beta_w(\theta_0) = \epsilon$$

(21)
$$\frac{d}{d\theta} \beta_w(\theta) |_{\theta=\theta_0} = m$$

for $w = w(k)$. Then $w(k)$ has power uniformly greater than or equal to that of any other test satisfying (20) and (21).

For $m = 0$ this becomes Neyman's theorem stating that the type A test is also of type A_1. The proof of the theorem however is independent of the value of

(23)
$$\frac{d}{d\theta} \beta_w(\theta) |_{\theta=\theta_0}$$

and hence carries over to arbitrary m.

II. If there exists any test satisfying (20) and (21) then there exists a number k for which $w(k)$ also satisfies (20) and (21).

To prove this let us determine, of all tests satisfying (20), the one which maximizes

$$(24) \qquad \frac{d}{d\theta} \beta_w(\theta) \big|_{\theta=\theta_0} = \int_w \frac{\partial}{\partial\theta} p_\theta(e) \big|_{\theta=\theta_0} de.$$

This can be done by means of the lemma of Neyman and Pearson [4, p. 11] which gives sufficient conditions for a region w, subject to restrictions

$$(25) \qquad \int_w f_i(e) \, de = a_i, \qquad\qquad (i = 1, \cdots, p),$$

to maximize an integral

$$(26) \qquad \int_w g(e) \, de.$$

According to this lemma the desired test is of the form

$$(25) \qquad \frac{\partial}{\partial\theta} p_\theta(e) \big|_{\theta=\theta_0} \geq a \cdot p_{\theta_0}(e)$$

provided a value of a exists for which this test satisfies (20). (25) is equivalent to

$$(26) \qquad P'(\theta_0) + T(e) \cdot Q'(\theta_0) \geq a \qquad\qquad \text{from (7)}$$

or, since $Q'(\theta_0) > 0$, to

$$(27) \qquad T(e) \geq b.$$

Thus, if a number b exists such that the test (27) satisfies (20), this test is the one maximizing (24). But such a number does exist, namely $f(-\infty)$. Therefore $w(-\infty)$ is the desired test.

Similarly it is easy to show that of all tests satisfying (20), $w(k_0)$ minimizes (24). But

$$(28) \qquad \frac{d}{d\theta} \beta_{w(k)}(\theta) \big|_{\theta=\theta_0} = \int_{T<k, T\geq f(k)} \frac{\partial}{\partial\theta} p_\theta(e) \big|_{\theta=\theta_0} de$$

is a continuous function of k, and therefore takes on all intermediate values, which establishes II.

III. From I. and II. we conclude that given any test w there exists a member of F which has power uniformly greater than or equal to that of w. For let w be any test of H. From the condition of regularity it follows that its power-function has a derivative at θ_0. By II. there exists a value of k such that the powerfunction of $w(k)$ has the same slope at θ_0, and from I. it follows that $w(k)$ is uniformly more powerful than w.

But from the lemma we see that none of the tests $w(k)$ is uniformly more powerful than any other. Hence all members of F are admissible, and the theorem is proved.

From the lemma and Theorem 1 we can conclude for all members of F the following optimum property:

COROLLARY 1: *Let* w *be any test, and let* w_0 *be any member of* F. *Then at least one of the two statements*

(29)
$$\beta_w(\theta) \leq \beta_{w_0}(\theta) \qquad \text{for all } \theta < \theta_0$$
$$\beta_w(\theta) \leq \beta_{w_0}(\theta) \qquad \text{for all } \theta > \theta_0$$

must hold.

The lemma and Theorem 1 also give the following result concerning most stringent tests, defined by Wald [1, p. 33].

COROLLARY 2: *There exists a uniformly most powerful of all most stringent tests. It is that unique member* w_0 *of* F *for which*

$$\underset{\theta < \theta_0}{\text{l.u.b.}} \left[\underset{w}{\text{l.u.b.}} \ \beta_w(\theta) - \beta_{w_0}(\theta) \right] = \underset{\theta > \theta_0}{\text{l.u.b.}} \left[\underset{w}{\text{l.u.b.}} \ \beta_w(\theta) - \beta_{w_0}(\theta) \right].$$

4. The effect on F of assumptions about the alternatives. Let us next consider how a restriction in the set of alternatives effects the family F. From the lemma it follows that there is no change as long as the set of alternatives contains values of θ both greater and less than θ_0. On the other hand, if the alternatives are restricted to values of θ greater than θ_0, say, the family F for testing H against these alternatives consists of only a single member, the test $w(-\infty)$, (and similarly for the other onesided case). This follows from

THEOREM 2: *Under conditions a. and b. the test* $w(-\infty)$ *is uniformly most powerful against the alternatives* $\theta > \theta_0$, *the test* $w(k_0)$ *is uniformly most powerful against the alternatives* $\theta < \theta_0$.

PROOF: Let w be any test. By Theorem 1 there exists a number k such that

(30)
$$\beta_w(\theta) \leq \beta_{w(k)}(\theta) \qquad \text{for all } \theta.$$

From the lemma it follows that

(31)
$$\beta_{w(k)}(\theta) \leq \beta_{w(k_0)}(\theta) \qquad \text{if } \theta < \theta_0$$
$$\beta_{w(k)}(\theta) \leq \beta_{w(-\infty)}(\theta) \qquad \text{if } \theta > \theta_0.$$

Combining (30) and (31) we have the desired result.

(It is also easy to prove Theorem 2 directly from the Neyman-Pearson lemma.)

In order to illustrate how the assumption of an a priori distribution of θ together with some information about this distribution affects F, let us consider a special case of the class of hypotheses discussed so far.

Let

(32)
$$p_\theta(x_1, \cdots, x_n) = c \cdot e^{-\frac{1}{2}\Sigma(x_i - \theta)^2}$$

so that $E = (X_1, X_2, \cdots, X_n)$ is a sample from a normal distribution with unit variance and unknown mean. We want to test the hypothesis

(33)
$$H: \theta = 0.$$

We shall show that if θ has a probability density function g which is symmetric

about the origin, then the family F for testing H consists, as might be expected, of a single member, the type A_1 test.

Our problem is to find the test w satisfying

$$(34) \qquad \int_w p_0(x_1, \cdots, x_n) \, dx_1 \cdots dx_n = \epsilon$$

and which maximizes

$$(35) \qquad \int_{-\infty}^{\infty} g(\theta) \int_w p_\theta(x_1, \cdots, x_n) \, dx_1 \cdots dx_n \cdot d\theta.$$

Inverting the order of integration, which is permissible in this case, the Neyman-Pearson lemma shows the desired test to be of the form

$$(36) \qquad \int_{-\infty}^{\infty} g(\theta) p_\theta(x_1, \cdots, x_n) \, d\theta \geq a \cdot p_0(x_1, \cdots, x_n)$$

provided a value of a exists for which (36) satisfies (34). Substituting from (32), (36) becomes

$$(37) \qquad f(\bar{x}) = \int_{-\infty}^{\infty} g(\theta) e^{-\frac{1}{2}\theta^2 + n\theta\bar{x}} \, d\theta \geq a$$

where

$$(38) \qquad \bar{x} = \frac{1}{n} \sum_{i=1}^{n} x_i.$$

Since

$$(39) \qquad \frac{d^2}{d\bar{x}^2} f(\bar{x}) > 0$$

the region (37) is either empty, which would contradict (34), or else can be described by inequalities

$$(40) \qquad \bar{x} \leq a_1, \qquad \bar{x} \geq a_2$$

where

$$(41) \qquad f(a_1) = f(a_2)$$

the latter equation becoming, on substitution from (37)

$$(42) \qquad \int_{-\infty}^{\infty} g(\theta) e^{-(n/2)\theta^2} (e^{na_2\theta} - e^{na_1\theta}) \, d\theta = 0.$$

If g is an even function, (42) is certainly satisfied when $a_1 = -a_2$. Our test then becomes

$$(43) \qquad \bar{x} \leq -a_2, \qquad \bar{x} \geq a_2$$

which for proper choice of a_2 satisfies (34) and is the well known type A_1 test.

5. Concluding remarks. Let us consider once more a probability density function satisfying a. and b. We have seen that the family F for testing H against the alternatives $\theta \neq \theta_0$ contains an infinity of elements unless we make some additional assumptions. On the other hand, if the principle of unbiasedness is accepted, F shrinks to a single element: the type A_1 test.

But unbiasedness does not insure power. Thus conceivably some other test might be more powerful than the test chosen, everywhere except in a small one sided neighbourhood of θ_0. That this is not so is shown by Corollary 1 to Theorem 1. This remark illustrates how intuitively appealing principles and a knowledge of the family F may be used in conjunction to arrive at a choice of a satisfactory test, when not enough information is available to make the choice compelling.

Finally, it should be pointed out that although we restricted our considerations to simple hypotheses, the notions developed also apply to composite hypotheses.

REFERENCES

[1] A. WALD, "On the principles of statistical inference". *Notre Dame Mathematical Lectures.* Number 1.

[2] J. NEYMAN, "L'estimation statistique traitée comme un problème classique de probabilité". *Conférences internationales des sciences mathématiques à Genève: Colloque d'Octobre 1937, sur le Calcul des Probabilités*, Paris, 1938.

[3] J. NEYMAN AND E. S. PEARSON, "Contributions to the theory of testing statistical hypotheses, Part II". *Statistical Research Memoirs*, Vol. II, London, 1938.

[4] J. NEYMAN AND E. S. PEARSON, "Contributions to the theory of testing statistical hypotheses, Part I". *Statistical Research Memoirs*, Vol. I, London, 1936.

Springer

Selected Works of E. L. Lehmann
Rojo, J. (Ed.)
2012, XLVIII, 1109 p., Hardcover
ISBN: 978-1-4614-1411-7