

## HSU'S WORK ON INFERENCE

BY E. L. LEHMANN

University of California, Berkeley

Hsu spent four years (1936-1940) at University College, London, where E. S. Pearson had recently succeeded his father in the chair of statistics and where, during the first two years, Neyman was a Reader in the Statistics Department. During this period Hsu wrote a remarkable series of papers on statistical inference which show the strong influence of the Neyman-Pearson point of view.

In 1938 Hsu's first two statistical papers appeared in Vol. II of the Neyman-Pearson edited *Statistical Research Memoirs*. The first of these [2] is concerned with what today is called the Behrens-Fisher problem. If  $X_i$  and  $Y_j$  ( $i = 1, \dots, m; j = 1, \dots, n$ ) denote samples from normal distributions  $N(\xi, \sigma^2)$  and  $N(\eta, \tau^2)$ , Hsu considers the class of statistics  $u = (\bar{Y} - \bar{X})^2 / (A_1 S_X^2 + A_2 S_Y^2)$  where  $S_X^2 = \sum (X_i - \bar{X})^2$  and  $S_Y^2 = \sum (Y_j - \bar{Y})^2$ . This reduces to  $u_1$ , Student's  $t$ , for  $A_1 = A_2 = N/mn(N - 2)$  where  $N = m + n$  and to the Behrens-Fisher statistic  $u_2$  for  $A_1 = 1/m(m - 1)$ ,  $A_2 = 1/n(n - 1)$ .

Hsu finds a series expansion for the density of  $u$ , and utilizes this to study the power function of the rejection regions  $u \geq C$  in terms of the parameters  $\theta = \tau^2/\sigma^2$  and  $\lambda = (\eta - \xi)^2 / (\frac{\sigma^2}{m} + \frac{\tau^2}{n})$ . It is an exact (not asymptotic) analysis, described by Scheffé (1970) as "a model of mathematical rigor". In the process, he obtains stochastic bounds for  $u_2$  which were later taken up independently and generalized by Hájek, Lawton and others (cf. Eaton and Olshen (1972)). Hsu's main conclusion, obtained by a combination of his analytical study with some numerical work, is that for  $\lambda = 0$  and varying  $\theta$  neither  $u_1$  nor  $u_2$  control the rejection probability at all well (except when  $m = n$ ) although of the two,  $u_2$  is less sensitive to variation of  $\theta$ .

In the second paper [3], Hsu treats the question of optimal estimators of the variance  $\sigma^2$  in the Gauss-Markov model. In the spirit of the Gauss-Markov theorem, he considers estimators  $Q$  which are (a) quadratic and (b) unbiased. In addition he imposes the restriction (c) that the variance of  $Q$  be independent of the unknown means. (This is a forerunner of the condition he imposed in [12] for the power function of analysis of variance tests).

Hsu then obtains a necessary and sufficient condition for the usual unbiased estimate  $S^2$  of  $\sigma^2$  to have uniformly minimum variance within this class of estimators. He illustrates the condition on a number of examples and, in particular, shows that  $S^2$  has the desired property in the one-sample case. The problem was

---

Received June 1978.

AMS 1970 subject classifications. Primary 01A70; secondary 62F05, 62F10.

Key words and phrases. Obituary, hypothesis testing, estimation.

taken up again by Rao (1952) who replaced (c) by the condition that  $Q$  be definite, while Seely (1971) instead considered an equivalent invariance condition. Hsu's paper may be regarded as the point of origin of the recent large literature on best quadratic estimates of variances and variance components.

The remaining papers in this series devoted to small-sample inference are concerned with the problem of testing univariate and multivariate linear hypotheses, particularly with power properties of the tests of these hypotheses. In the first of these papers [4], Hsu obtains the power of Hotelling's  $T^2$ -test, showing that the nonnull distribution of  $T^2$  is that of the ratio of a noncentral  $\chi^2$  to an independent central  $\chi^2$  variable. He also proves that the test is in a certain sense locally most powerful. In addition, he points out some new applications of  $T^2$ .

From a study of  $T^2$  it was a natural step to a multivariate extension of the general linear (univariate) hypothesis discussed by Kolodziejczyk (1935). Generalizing some examples considered earlier in the literature, Hsu in [9] formulates the general multivariate linear hypothesis in its canonical form. He obtains the nonnull distribution of the likelihood ratio test statistic when the covariance matrix is known. For the case of unknown covariance matrix, if  $\theta_i$  denote the nonvanishing roots of the associated determinantal equation, Hsu considers the test statistics  $W = \Pi(1 - \theta_i)$  and  $V = \Sigma \theta_i / (1 - \theta_i)$ , and proves that their asymptotic power agrees as the sample size tends to infinity. In a succeeding paper [13], Hsu shows how the general multivariate regression problem can be reduced to the earlier canonical form so that the theory of [9] becomes applicable.

Perhaps the most important paper in this series is [12], where Hsu obtains the first optimum property for the likelihood ratio test of the univariate linear hypothesis, in fact essentially the first nonlocal optimum property for any hypothesis specifying the value of more than one parameter. It had been shown by Kolodziejczyk that a uniformly most powerful similar test does not exist for linear hypotheses with more than one constraint. If  $\lambda$  denotes the noncentrality parameter, which is the only parameter on which the nonnull distribution of the likelihood ratio test for this problem depends, Hsu proves that this test is uniformly most powerful among all tests with power function depending only on  $\lambda$ . It was later realized that this condition is equivalent to invariance of the power function under the natural class of transformations. The resulting optimum property is essentially also equivalent to the fact that the test is uniformly most powerful invariant.

This paper started two lines of development. On the one hand, Hsu's formulation was applied to multivariate problems (Hotelling's  $T^2$  and the multiple correlation coefficient) by his student Simaika (1941) and modifications of the formulation were considered by Wald (1942), Hsu (1945) [19], Wolfowitz (1949), Lehmann (1959) and others. On the other hand, in the course of this paper Hsu provided a new method for obtaining all similar tests. At Hsu's suggestion this method was applied to other problems by Simaika (1941) and Lehmann (1947), and was formalized by means of the concept of completeness by Lehmann and Scheffé (1950).

## REFERENCES

- EATON, MORRIS L. and OLSHEN, RICHARD A. (1972). Random quotients and the Behrens-Fisher problem. *Ann. Math. Statist.* **43** 1852–1860.
- HSU, P. L. (1945). On the power functions for the  $E^2$ -test and the  $T^2$ -test. *Ann. Math. Statist.* **16** 278–286.
- KOŁODZIECZYK, S. (1935). On an important class of statistical hypotheses. *Biometrika* **21** 161–190.
- LEHMANN, E. L. (1947). On optimum tests of composite hypotheses with one constraint. *Ann. Math. Statist.* **18** 473–494.
- LEHMANN, E. L. (1959). Optimum invariant tests. *Ann. Math. Statist.* **30** 881–884.
- LEHMANN, E. L. and SCHEFFÉ, HENRY (1950). Completeness, similar regions and unbiased estimation, *Sankhyā* **10** 305–340.
- RAO, C. RADHAKRISHNA (1952). Some theorems on minimum variance estimation. *Sankhyā* **12** 27–44.
- SCHEFFÉ, HENRY (1970). Practical solutions of the Behrens-Fisher problem. *J. Amer. Statist. Assoc.* **65** 1501–1508.
- SEELY, JUSTUS (1971). Quadratic subspaces and completeness. *Ann. Math. Statist.* **42** 710–721.
- SIMAİKA, J. B. (1941). On an optimum property of two important statistical tests. *Biometrika* **32** 70–80.
- WALD, ABRAHAM (1942). On the power function of the analysis of variance test. *Ann. Math. Statist.* **13** 434–439.
- WOLFOWITZ, J. (1949). The power of the classical tests associated with the normal distribution. *Ann. Math. Statist.* **20** 540–551.

DEPARTMENT OF STATISTICS  
UNIVERSITY OF CALIFORNIA  
BERKELEY, CALIFORNIA 94720