AN INEQUALITY CONCERNING TESTS OF FIT OF THE KOLMOGOROV-SMIRNOV TYPE

By Georges Vandewiele and Marc Noé

Free University of Brussels and M.B.L.E. Research Laboratory, Brussels

1. Introduction. Let $F_n(x)$ be the sumpolygon (empirical distribution function) of a sample of size n from a continuous distribution function $F(x)$. Let $K(x)$, $G_1(x)$, $G_2(x)$, $H_1(x)$, $H_2(x)$ be functions of x, such that for all x,

$$G_1(x) \geqq G_2(x); \qquad H_2(x) \geqq H_1(x).$$

The object of the present paper is to prove the following inequalities

$$(1) \quad P[\inf_x (F_n - K) \geqq 0 \mid \inf_x (G_1 - F_n) \geqq 0, \quad \inf_x (F_n - H_2) \geqq 0] \geqq P[\inf_x (F_n - K) \geqq 0 \mid \inf_x (G_2 - F_n) \geqq 0, \quad \inf_x (F_n - H_1) \geqq 0],$$

$$(2) \quad P[\inf_x (K - F_n) \geqq 0 \mid \inf_x (G_1 - F_n) \geqq 0, \quad \inf_x (F_n - H_2) \geqq 0] \leqq P[\inf_x (K - F_n) \geqq 0 \mid \inf_x (G_2 - F_n) \geqq 0, \quad \inf_x (F_n - H_1) \geqq 0],$$

where all probabilities are supposed to exist. Since these inequalities are symmetrical, it suffices to prove one of them.

These inequalities provide an approximation for the distribution of two-sided statistics of the Kolmogorov-Smirnov type. Such a distribution is written

$$P\{\sup_x n^{\frac{1}{2}}|F_n(x) - F(x)|\psi[F(x)] \leqq \lambda\}$$

or more generally

$$(3) \qquad P_n = P[\inf_x (G_2 - F_n) \geqq 0, \quad \inf_x (F_n - H_2) \geqq 0].$$

In order to approximate P_n, take $H_1(x)$ and $H_2(x)$ in (1) smaller than zero for all x and replace $K(x)$ in (1) by $H_2(x)$; similarly take $G_1(x)$ and $G_2(x)$ in (2) larger than 1 for all x and replace $K(x)$ in (2) by $G_1(x)$. One then easily obtains the upper bound

$$(4) \quad P_n \leqq P_n' \, P[\inf_x (G_2 - F_n) \geqq 0] \cdot P[\inf_x (F_n - H_2) \geqq 0] \cdot \{P[\inf_x (G_1 - F_n) \geqq 0] \cdot P[\inf_x (F_n - H_1) \geqq 0]\}^{-1}$$

where $P_n' = P[\inf_x (G_1 - F_n) \geqq 0, \inf_x (F_n - H_1) \geqq 0]$. If now G_1 and H_1 are chosen close to G_2 resp. H_2, but such that P_n' is more easily calculable than P_n, then (4) provides an interesting approximation of (3). A lower bound can be found in a similar way.

Wald and Wolfowitz [3] and [4] have given the following two bounds for P_n

$$(5) \qquad P_n \leqq P[\inf_x (G_2 - F_n) \geqq 0] \cdot P[\inf_x (F_n - H_2) \geqq 0],$$

$$P_n \geqq P[\inf_x (G_2 - F_n) \geqq 0] + P[\inf_x (F_n - H_2) \geqq 0] - 1.$$

Received 27 December 1966; revised 23 February 1967.