The Annals of Statistics
1976, Vol. 4, No. 1, 214–215

ON THE ASYMPTOTIC NORMALITY OF KENDALL'S RANK CORRELATION STATISTIC

BY MILOSLAV JIRINA

The Flinders University of South Australia

Kendall's rank correlation statistic $T_n = \sum_{i>j} \operatorname{sgn}(X_i - X_j) \cdot \operatorname{sgn}(Y_i - Y_j)$ is well known to be asymptotically normally distributed under the null hypothesis of independence as the sample size $n \to \infty$. In the note it is shown that this assertion can be obtained easily from the recurrence formula $p_n(t) = (1/n) \sum_{k=1}^{n} p_{n-1}(t - 2k + n + 1)$ for the probability distribution p_n of T_n (see Kendall (1970), e.g.). This recurrence formula implies that T_n has the same distribusion as a sum of $(n - 1)$ well defined independent random variables to which the Lyapunov criterion applies.

Kendall's rank correlation test is based on the statistic $T_n = \sum_{i>j} \operatorname{sgn}(X_i - X_j) \operatorname{sgn}(Y_i - Y_j)$, where $(X_1, Y_1), \cdots, (X_n, Y_n)$ is a sample of size n from a continuous bivariate population and $\operatorname{sgn} y = 1$ if $y > 0$, $\operatorname{sgn} y = -1$ if $y < 0$. It is well known that

(1) *under the null hypothesis H_0 of independence, the statistics T_n are asymptotically normally distributed as $n \to \infty$.*

Kendall uses the moment method in proving (1) (see [1], 5.8 e.g.). The aim of this note is to show that (1) can be obtained easily by proving that T_n has under H_0 the same distribution as a sum of $n - 1$ independent random variables. A similar method was used in [2], however the proof is much longer than the present one.

PROOF OF (1). For each integer t and for each $n \geqq 2$ put $p_n(t) = \operatorname{Prob}(T_n = t \mid H_0)$. The following recurrence formula for p_n will be used (see [1], 5.2 e.g.):

$$p_n(t) = \frac{1}{n} \sum_{k=1}^{n} p_{n-1}(t - 2k + n + 1) \tag{2}$$

for any integer t and $n \geqq 3$. For each integer t and each $n \geqq 2$ put

$$\begin{aligned} q_n(t) &= \frac{1}{n} \quad \text{if} \quad t = -n+1, -n+3, \cdots, n-3, n-1 \\ &= 0 \quad \text{otherwise.} \end{aligned}$$

Using the convolution operation $*$ we can rewrite (2) in the form $p_n = q_n * p_{n-1}$ which, together with $p_2 = q_2$, yields $p_n = q_n * q_{n-1} * \cdots * q_2$. Hence under H_0,

---

Received January 1975; revised April 1975.

AMS 1970 *subject classification*. Primary 62E20.

Key words and phrases. Kendall rank correlation statistic, asymptotic normality, recurrence formula.

T_n has the same distribution as the sum $\sum_{k=2}^n Z_k$, where $Z_2, \cdots, Z_n$ are independent integer-valued random variables such that Z_k has probability distribution q_k. Clearly, $E(Z_k) = 0$, $E(Z_k^2) = \frac{1}{3}(k^2 - 1)$, $E(|Z_k|^3) \leqq k^3$ and $\sum_{k=2}^n E(|Z_k|^3)/[\sum_{k=2}^n E(Z_k^2)]^{\frac{3}{2}} \leqq cn^{-\frac{1}{2}} \to_n 0$. Hence, by the Lyapunov theorem, T_n are asymptotically normal.

REFERENCES

[1] KENDALL, M. G. (1970). *Rank Correlation Methods*, 4th ed. Griffin, London.

[2] TERPSTRA, T. J. (1952). The asymptotic normality and consistency of Kendall's test against trend, when ties are present in one ranking. *Indag. Math.* **14** 327–333.

SCHOOL OF MATHEMATICAL SCIENCES
THE FLINDERS UNIVERSITY OF SOUTH AUSTRALIA
BEDFORD PARK, SOUTH AUSTRALIA 5042