Časopis pro pěstování matematiky a fysiky

Antonín Špaček

Note on successive cumulative sums of independent random variables

Časopis pro pěstování matematiky a fysiky, Vol. 74 (1949), No. 1, 41--45

Persistent URL: http://dml.cz/dmlcz/109148

NOTE ON SUCCESSIVE CUMULATIVE SUMS OF INDEPENDENT RANDOM VARIABLES.

ANTONÍN ŠPAČEK, Praha.

(Received April 4, 1949.)

Let z be a random variable and $z_1, z_2, z_3, \ldots$ a sequence of independent random variables, each with the same distribution as z. We shall always suppose, that the distribution of z has positive variance $D^2(z)$, which excludes trivial cases. Define $Z_m = z_1 + z_2 + \ldots + z_m$ for $m = 1, 2, 3, \ldots$ and choose three constants $b < 0 < a$, $d > 0$. Then $v_m = z_m + d$ for $m = 1, 2, 3, \ldots$ also are independent and identically distributed random variables according to the distribution of $v = z + d$. Let $V_m = v_1 + v_2 + \ldots + v_m$ for $m = 1, 2, 3, \ldots$. Then $V_m = Z_m + md$ for $m = 1, 2, 3, \ldots$. Define the random variable n as the smallest positive integer, for which $V_n \leqq b$, or $Z_n \leqq b$, $V_n \geqq a$, or $Z_n \geqq a$, i. e. $Z_n \leqq b - nd$, or $a - nd \leqq Z_n \leqq b$, or $a \leqq Z_n$.

This generalization of WALDS' theory is of particular interest in the following sampling inspection scheme: Let $x_1, x_2, x_3, \ldots$ be a sequence of independent observations on the random variable x, which admits a discrete distribution or probability density function $f(x, \vartheta)$ depending upon the single unknown parameter ϑ. For particular values $\vartheta_1 > \vartheta_0 \geqq \vartheta_0' > \vartheta_1'$ of ϑ, define the random variables $z = \log[f(x, \vartheta_1)/f(x, \vartheta_0)]$, $v = \log[f(x, \vartheta_1')/f(x, \vartheta_0')]$, each of which, evidently, is a function of the other. For a special class of discrete distributions and densities, this function is linear, for example the three most important distributions, the binomial, the POISSON and the normal distributions belong to this class. In the case of POISSON and normal distributions, if ϑ is the mean, we have the special linear function $v = z + d$ for $\vartheta_1 - \vartheta_0 = \vartheta_1' - \vartheta_0'$ and if $f(x, p)$ is a binomial distribution and $1 > p_1 > p_0 \geqq p_1' > p_0' > 0$, then $v = z + d$ holds for $p_1' = 1 - p_0$ and $p_0' = 1 - p_1$. The inspection procedure is defined as follows: We take observations $x_1, x_2, x_3, \ldots$ and compute the corresponding values $z_1, z_2, z_3, \ldots$ as long as $b < Z_m < a$ or $b < Z_m + md < a$ is satisfied. The first time that neither Z_m nor $Z_m + md$ lies in the open interval (b, a), inspection is terminated. At the termination of this procedure the lot is accepted, if $a - md \leqq Z_m \leqq b$, and rejected otherwise.

The purpose of this note is to show that all theorems of WALD'S theory [1] remain true for our more general case. In particular, we shall show, that the following theorems hold:

(i) *The random variable n has finite moments of all orders.*

(ii) *The fundamental identity* (15) *in* [1] *holds.*

(iii) *The differentiability of the fundamental identity, as expressed by theorem* 2,1 *in* [2] *holds.*

The proof will result from the following two lemmas:

(iv) *There exists a positive number c and a positive number $p < 1$, such that $P(n > k) < cp^k$ for $k = 1, 2, 3, \ldots$.*

(v) *The conditional expectation $E(e^{Z_n t} \mid n = m)$ is bounded independently of m for each fixed real t.*

Proof of lemma (iv): Let us denote by $I_\varkappa$ the open interval $(b - a + \varkappa d,\ a - b + \varkappa d)$ for $\varkappa = 0, \pm 1, \pm 2, \ldots$ and define $Z_0 = 0$; $Z_m^{(j)} = z_{mj+1} + z_{mj+2} + \ldots + z_{mj+m}$ for $m = 1, 2, 3, \ldots,\ j = 0, 1, 2, 3, \ldots$. Then

$$Z_{mj} = Z_{m(j-1)} + Z_m^{(j-1)} \tag{1}$$

for $j = 1, 2, 3, \ldots$ and we shall prove the following proposition (π): *If $n > k \geqq m$, then*

$$Z_m^{(j-1)} \in I_{-mj} + I_{-m} + I_0 + I_{m(j-1)}, \tag{2}$$

for each positive integer j, which satisfies the condition $1 \leqq j \leqq k/m$. Suppose that this is not so and denote by j_0 the least subscript which *satisfies the condition $1 \leqq j_0 \leqq k/m$ and for which* (2) does not hold. If $j_0 = 1$, then by (2) $Z_m^{(0)} = Z_m$ non $\in I_{-m} + I_0$ and hence Z_m does not lie in the sum of the open intervals $(b - md,\ a - md)$ and (b, a), i. e. $n \leqq m \leqq k$ contrary to hypothesis. Let $j_0 > 1$. Then either $n \leqq m(j_0 - 1) < mj_0 \leqq m(k/m) = k$, which is by hypothesis impossible, or $n > m(j_0 - 1)$. If the last inequality holds, then $Z_{m(j_0-1)}$ lies in the sum of the open intervals $(b - m(j_0 - 1)d,\ a - m(j_0 - 1)d)$ and (b, a), by (2) and by the definition of j_0, we have $Z_m^{(j_0-1)}$ non $\in I_{-mj_0} + I_{-m} + I_0 + I_{m(j_0-1)}$ and therefore by (1), Z_{mj_0} does not lie in the sum of the open intervals $(b - mj_0 d,\ a - mj_0 d)$ and (b, a), i. e. $n \leqq mj_0 \leqq m(k/m) = k$ contrary to hypothesis. The proposition (π) is thus proved. Since $z_1, z_2, z_3, \ldots$ are independent and identically distributed, the same holds also for $Z_m^{(0)}, Z_m^{(1)}, Z_m^{(2)}, \ldots$, such that for $k \geqq m$

$$P(n > k) \leqq \prod_{j=1}^{j_k} P(Z_m \in I_{-m} + I_{-mj} + I_0 + I_{m(j-1)}), \tag{3}$$

where j_k is the greatest positive integer $\leqq k/m$. We shall now establish the following proposition (ϱ): *There exists a pair of positive integers r and s and a positive number $q < 1$, such that $P(Z_r \in I_{-rj} + I_{-r} + I_0 + I_{r(j-1)}) < q$*

for each subscript $j > s$. If (ϱ) is true, then by (3) for $k \geqq r(s+1)$

$$P(n > k) \leqq \prod_{j=s+1}^{j_k} P(Z_r \in I_{-rj} + I_{-r} + I_0 + I_{r(j-1)}) < q^{j_k - s} < q^{\frac{k}{r} - \frac{s+1}{r}}$$

and if we substitute c for $q^{-\frac{s+1}{r}}$ and p for $q^{\frac{1}{r}}$, lemma (iv) is proved under the assumption that (ϱ) holds. We shall show firstly, that *there exists a positive integer* m_0 *such that*

$$P(Z_m \in I_{-m} + I_0) < 1 \tag{4}$$

for each subscript $m > m_0$. We distinguish three cases:

$$0 \neq E(z) \neq -d, \ E(z) = 0, \ E(z) = -d.$$

Let $0 \neq E(z) \neq -d$ and suppose, that to each positive integer m there exists a subscript $k_m > m$ such that

$$P(Z_{k_m} \in I_{-k_m} + I_0) = 1, \tag{5}$$

for $m = 1, 2, 3, \ldots$, i. e. there exists an infinite sequence of positive integers $k_1 < k_2 < k_3 < \ldots$, such that (5) holds for $m = 1, 2, 3, \ldots$. Then

$$P[(a-b)^2 - dk_m(a-b) < Z_{k_m}^2 + dk_m Z_{k_m} < (a-b)^2 + \\ + dk_m(a-b)] = 1$$

for $m = 1, 2, 3, \ldots$, and since

$$E(Z_{k_m}^2 + dk_m Z_{k_m}) = k_m D^2(z) + k_m^2 E(z)[E(z) + d],$$

we have

$$-d(a-b) \leqq D^2(z) + k_m E(z)[E(z) + d] \leqq (a-b)^2 + d(a-b)$$

for $m = 1, 2, 3, \ldots$, which is impossible. Let now $E(z) = 0$. For a properly choosen $\delta > 0$, we have $P(z > \delta) > 0$. Suppose that $P(z > \delta) = 0$ for each $\delta > 0$, hence $P(z > 0) = 0$, i. e. $P(z \leqq 0) = 1$. Then either $P(z = 0) = 1$, which is impossible, because $D^2(z) > 0$, or $P(z \leqq -\omega) > 0$ for a propperly choosen $\omega > 0$, because otherwise $P(z \leqq 0) = 0 < 1$, hence $E(z) \leqq -\omega P(z | z \leqq -\omega) < 0$ contrary to hypothesis. Since the random variables $z_1, z_2, z_3, \ldots$ are independent, each with the same distribution as z, and $z_k > \delta$ for $k = 1, 2, 3, \ldots, m$ implies $Z_m > a - b$ for $m > m_0 \geqq (a-b)/\delta$, hence $0 < [P(z > \delta)]^m \leqq P(Z_m > a - b) \leqq \leqq P(Z_m \geqq a - b)$, i. e. $P(Z_m < a - b) < 1$ and the last inequality implies (4) for each subscript $m > m_0$. Similarly, if $E(z) = -d$, then $E(v) = 0$ and we have $P(V_m > b - a) < 1$ for a sufficiently large m, i. e. $P(Z_m > b - a - md) < 1$, hence (4) holds for each subscript $m > m_0$. If k_0 is a positive integer $\geqq 2(a-b)/d$ and $r = \max(m_0 + 1, k_0)$, we have

$$P(Z_r \in I_{-r} + I_0) < 1, \tag{6}$$

$$I_{jr} \cdot I_{kr} = 0, \ j \neq k, \ j = 0, \pm 1, \pm 2, \ldots, \ k = 0, \pm 1, \pm 2, \ldots. \tag{7}$$

We shall further show that *to each* $\eta > 0$ *there exists a positive integer* ν_0 *such that*

$$P(Z_r \in I_{-rj} + I_{r(j-1)}) < \eta \tag{8}$$

for each subscript $j > \nu_0$. Suppose that a particular $\eta = \eta_0 > 0$ has the property that, to each positive integer j, there exists a subscript $m_j > j$ such that

$$P(Z_r \in I_{-rm_j} + I_{r(m_j-1)}) \geqq \eta_0, \tag{9}$$

i. e. there exists an infinite sequence of positive integers $m_1 < m_2 < < m_3 < \dots$, such that (9) holds for $j = 1, 2, 3, \dots$. But this is not possible, because, for a positive integer $j_0 > 1/\eta_0$ we have by (9) and (7)

$$P(Z_r \in \sum_{j=1}^{j_0}(I_{-rm_j} + I_{r(m_j-1)})] = \sum_{j=1}^{j_0} P(Z_r \in I_{-rm_j} + I_{r(m_j-1)}) \geqq j_0\eta_0 > 1,$$

and therefore (8) holds for each subscript $j > \nu_0$. Let now

$$\eta = \tfrac{1}{2}[1 - P(Z_r \in I_{-r} + I_0)].$$

By (6), $\eta > 0$ and by (8), there exists a positive integer s, such that

$$P(Z_r \in I_{-rj} + I_{r(j-1)}) < \tfrac{1}{2}[1 - P(Z_r \in I_{-r} + I_0)] = \\ = \tfrac{1}{2}[1 + P(Z_r \in I_{-r} + I_0)] - P(Z_r \in I_{-r} + I_0),$$

i. e., by (7)

$$P(Z_r \in I_{-rj} + I_{-r} + I_0 + I_{r(j-1)}) < \tfrac{1}{2}[1 + P(Z_r \in I_{-r} + I_0)] < 1,$$

for each subscript $j > s$. If we substitute q for $\tfrac{1}{2}[1 + P(Z_r \in I_{-r} + I_0)]$, then ($\varrho$) is proved and the proof of lemma (iv) is complete.

Proof of lemma (v): Since

$$E(e^{Z_n t} \mid n = m) = \\ = P(Z_n \leqq b - nd \mid n = m)\, E(e^{Z_n t} \mid Z_n \leqq b - nd,\, n = m) + \\ + P(a - nd \leqq Z_n \leqq b \mid n = m)\, E(e^{Z_n t} \mid a - nd \leqq Z_n \leqq b,\, n = m) + \\ + P(Z_n \geqq a \mid n = m)\, E(e^{Z_n t} \mid Z_n \geqq a,\, n = m),$$

we have for $t > 0$

$$E(e^{Z_n t} \mid n = m) \leqq 2e^{bt} + P(Z_n \geqq a \mid n = m)\, E(e^{Z_n t} \mid Z_n \geqq a,\, n = m) = \\ = 2e^{bt} + e^{at}e^{-xt}\, P(z \geqq x)\, E(e^{zt} \mid z \geqq x) = 2e^{bt} + e^{at}e^{-xt}\int_{z \geqq x} e^{zt}\, \mathrm{d}F(z) \leqq \\ \leqq 2e^{bt} + e^{at}\,\varphi(t),$$

where $\varphi(t)$ is the moment generating function; similarly for $t < 0$

$$E(e^{Z_n t} \mid n = m) \leqq 2e^{at} + e^{bt}\,\varphi(t)$$

and finally $E(e^{Z_n t} \mid n = m) = 1$ for $t = 0$, which completes the proof.

Proof of theorem (i): For each positive integer r the moment of r-th order is $E(n^r) = \sum_{k=1}^{\infty} k^r\, P(n = k)$. Since clearly $P(n = k) \leqq P(n >$

$> k - 1)$ and by lemma (iv) $E(n^r) < c \sum_{k=1}^{\infty} k^r p^{k-1}$ for $k = 1, 2, 3, \ldots$, hence theorem (i) follows at once from the D'ALEMBERT'S criterion.

The proof of theorem (ii) is the same as in [1] and it is enough to note only, that $|E(e^{Z_m t} | n > m)| \leq e^{a|t|}$ and by lemma (iv) $\lim_{m \to \infty} P(n > m) = 0$.

To prove theorem (iii), it is sufficient to show that the expectation 2,16 in [2] is finite and this follows at once from lemma (v) and theorem (i)

By differentiating the fundamental identity at $t = 0$, we obtain the well known formulae

$$E(n) = \frac{E(Z_n)}{E(z)} \text{ for } E(z) \neq 0, \tag{10}$$

$$E(n) = \frac{E(Z_n^2)}{E(z^2)} \text{ for } E(z) = 0. \tag{11}$$

It is easy to show, that the above mentioned theorems remain true for $d < 0$ and therefore it is not necessary to use the formula (11), because if $E(z) = 0$, then $E(v) = d$ and we have $E(n) = E(V_n)/E(v)$. Under the assumption that the coditions of lemma 2 in [1] hold (respectively without the condition $E(z) \neq 0$), a formula may be obtained from the fundamental identity for $P(a - nd \leq Z_n \leq b)$, which is analogous to the formula (18) in [1].

Tesla National Corporation, Praha.

REFERENCES.

[1] A. WALD: On cumulative sums of random variables, Ann. Math. Stat., vol. 15 (1944), pp. 283—296.
[2] A. WALD: Differentiation under the expectation sign in the fundamental identity of sequential analysis, Ann. Math. Stat., vol. 17 (1946), pp. 493—497.

O postupných kumulativních součtech nezávislých náhodných proměnných.

(Obsah předešlého článku.)

Jedná se o jisté zobecnění WALDOVY theorie postupných kumulativních součtů nezávislých náhodných proměnných, kterého lze použít při testování hypotéz postupnými výběry. Dokazuje se, že při takovém zobecnění zůstávají v platnosti všechny věty WALDOVY theorie.