An estimate of waiting times in the problem of covering the unit interval

Silvia Bontaş

Abstract

T_ε represents the coverage time of unit interval when it is bombed with segments of 2ε length. A formula and estimates of the expectation of T_ε are obtained.

Mathematics Subject Classification: 60E15.
Key words: expectation, random variable, probability field, covering, measurability.

1 Introduction

The statement of the problem. The unit interval $I = [0, 1]$ is randomly bombed with segments of length 2ε. The following problem arises: how long does it take to cover with such segments the whole interval $[0, 1]$? In other words, how many random segments of length 2ε are necessary to cover the whole unit interval ?

Let (ω, K, P) be a probability field and $(U_n)_{n\geq 1}$ a sequence of independent random variables, uniformly distributed on I, defined on this probability field. We define the minimal waiting time until the complete covering of the interval I with segments of length 2ε centered at $U_1, \ldots, U_n$ by

$$(1.1)\qquad T_\varepsilon(\omega) = \inf \left\{ n \geq 1 \;\middle|\; I \subset \bigcup_{k=1}^{n} [U_k(\omega) - \varepsilon, U_k(\omega) + \varepsilon] \right\}.$$

In other words, $T_\varepsilon(\omega)$ represents the first moment at which the segment I was entirely destroyed due to the bombing with segments of length 2ε. For any $a \in I$ we define the minimal expectance time until attending the point a as consequence of covering the interval I with segments of length 2ε by

$$(1.2)\qquad T_{\varepsilon,a}(\omega) = \inf \left\{ n \geq 1 \;\middle|\; a \in \bigcup_{k=1}^{n} [U_k(\omega) - \varepsilon, U_k(\omega) + \varepsilon] \right\},$$

Proceedings of The 3-rd International Colloquium "Mathematics in Engineering and Numerical Physics" October 7-9 , 2004, Bucharest, Romania, pp. 40-50.

which means that $T_{\varepsilon,a}(\omega)$ is the first moment at which the point $a \in I$ was attended as consequence of the bombing with segments of length 2ε. Obviously we have

$$T_\varepsilon = \sup_{a\in I} T_{\varepsilon,a}. \tag{1.3}$$

From (1.3) it does not follow that T_ε is a random variable, since the supremum of a family of measurable functions is not necessary a measurable function. We shall further consider another approach for T_ε, which shall point out its measurability.

Let $C_n(\omega) = \{U_1(\omega), U_2(\omega), \dots, U_n(\omega)\}$. We remark that

$$a \in \bigcup_{k=1}^{n} [U_k(\omega) - \varepsilon, U_k(\omega) + \varepsilon] \Leftrightarrow d(a, C_n(\omega)) \le \varepsilon, \tag{1.4}$$

where for any set $C \subseteq I$,

$$d(a, C) = \inf\{|x - a| \mid x \in C\}. \tag{1.5}$$

The Hausdorff distance between two closed sets $A, B \subseteq \mathbb{R}$ is

$$D(A, B) = \sup_{a\in A} d(a, B) + \sup_{b\in B} d(A, b). \tag{1.6}$$

In particular, if $A \subseteq B$, then

$$A(A, B) = \sup_{b\in B} d(A, b). \tag{1.7}$$

Having in view these relations, we can write

$$T_{\varepsilon,a}(\omega) = \inf\{n \ge 1 \mid d(a, C_n(\omega)) \le \varepsilon\} \tag{1.8}$$

and

$$T_\varepsilon(\omega) = \inf\{n \ge 1 \mid d(I, C_n(\omega)) \le \varepsilon\}. \tag{1.9}$$

Since for any $A \subseteq I$ we have $D(I, A) = \sup_{x\in\Gamma} d(x, A)$ (where Γ is a countable subset of I, dense in I), it follows that T_ε is measurable.

2 The distribution of T_ε

Let $n \ge 2$ be fixed and let $V = (U_{(1)}, U_{(2)}, \dots, U_{(n)})$ be the order statistics of the vector $U = (U_1, U_2, \dots, U_n)$. In other words, the components of V are the components of U, permuted such that $U_{(1)} \le U_{(2)} \le \dots \le U_{(n)}$. Consider the set $C_n = \{U_{(1)}, U_{(2)}, \dots, U_{(n)}\}$ defined above. Then

$$\begin{aligned} d(a, C_n) = {} & \min_{1\le k\le n} |a - U_{(k)}| = |a - U_{(1)}| 1_{\left[0, \frac{U_{(1)}+U_{(2)}}{2}\right]}(a) + \\ & + \sum_{k=1}^{n-1} |a - U_{(k)}| 1_{\left[\frac{U_{(k-1)}+U_{(k)}}{2} + \frac{U_{(k)}+U_{(k+1)}}{2}\right]}(a) + \\ & + |a - U_{(n)}| 1_{\left[\frac{U_{(n-1)}+U_{(n)}}{2}, 1\right]}(a). \end{aligned}$$

According to these relations, we have

$$d(I,C_n)=\max\left(U_{(1)},\frac{U_{(2)}-U_{(1)}}{2},\ldots,\frac{U_{(n)}-U_{(n-1)}}{2},1-U_{(n)}\right), \tag{2.1}$$

and hence, from (1.9) we get

$$\begin{aligned}\{T_\varepsilon\le n\}=\{U_{(1)}\le\varepsilon,\quad &U_{(2)}-U_{(1)}\le 2\varepsilon,\ldots,U_{(n)}-U_{(n-1)}\le 2\varepsilon,\\ &U_{(n)}\ge 1-\varepsilon\}.\end{aligned} \tag{2.2}$$

For computing $F_\varepsilon(n)=P(T_\varepsilon\le n)$ we use that the distribution of V is

$$P\circ V^{-1}=(n!\ 1_\varepsilon)\lambda^n, \tag{2.3}$$

where λ^n is the Lebesgue measure in $\mathbb{R}^n$, and $E=\{x\in\mathbb{R}^n\mid 0\le x_1\le x_2\le\ldots\le x_n\le 1\}$ (see Wilks [5, p.236]). We perform the change of coordinates

$$y=f(v): y_1=v_1, y_2=v_2-v_1,\ldots,y_n=v_n-v_{n-1}, \tag{2.4}$$

whose Jacobian is equal to 1. Let $Y=f(V)$. Then

$$P\circ Y^{-1}=(n!\ 1_S)\lambda_n, \tag{2.5}$$

where $S=S_n=\operatorname{co}(\{0,e_1,\ldots,e_n\})$ (we have denoted by 0 the null vector of $\mathbb{R}^n$, and $(e_j)_{1\le j\le n}$ is the canonic basis of $\mathbb{R}^n$). Then $S=\{x\in[0,1]^n\mid x_1+x_2+\ldots+x_n\le 1\}$ (see Wilks [5, p. 237]). We denote this distribution with μ. We conclude that the distribution of T_ε is given by:

Proposition 2.1. *Assuming that the previous conditions are fulfilled, the following relations hold true:*

$$\begin{aligned}P(T_\varepsilon\le n)=P(Y_1\le\varepsilon,\quad &Y_2\le 2\varepsilon,\ldots,Y_n\le 2\varepsilon,\\ &Y_1+\ldots+Y_n\ge 1-\varepsilon)=\mu(A_n),\end{aligned} \tag{2.6}$$

where

$$A_n=\{x\in S\mid x_1\le\varepsilon,x_2\le 2\varepsilon,\ldots,x_n\le 2\varepsilon,x_1+\ldots+x_n\ge 1-\varepsilon\}. \tag{2.7}$$

We shall determine a formula for computing $\mu(A_n)$. To this aim we shall use two results from [1, pp. 43-44].

Lemma 2.2. *Let $S_j(\varepsilon)=\varepsilon e_j+(1-\varepsilon)S$ the homotheties of S. Then the closure of $S\backslash A_n$ is*

$$\overline{S\backslash A_n}=S_0(\varepsilon)\cup S_1(\varepsilon)\cup\bigcup_{j=2}^{n}S_j(2\varepsilon). \tag{2.8}$$

Lemma 2.3. *Let $k\le n$ be fixed, let $\varepsilon_t>0, t=\overline{1,n}$ and let $\varepsilon=\varepsilon_1+\ldots+\varepsilon_k$. We denote $J=\{0\le j_1<j_2<\ldots<j_k\le n\}\subseteq\{0,1,\ldots,n\}$. Then*

$$\bigcap_{t=1}^{n}S_{j_t}(\varepsilon_t)=\begin{cases}\displaystyle\sum_{i=1}^{n}\varepsilon_t e_{j_t}+(1-\varepsilon)S, & \text{for } \varepsilon\le 1\\ \varnothing, & \text{for } \varepsilon>1.\end{cases} \tag{2.9}$$

We can compute now the probability that the expectance time T_ε be greater than a given n.

Proposition 2.4. *Let $0 < \varepsilon \leq 1$. Then*

$$P(T_\varepsilon > 1) = \min(1, 2(1-\varepsilon)). \tag{2.10}$$

For any $n \geq 2$, we have

$$P(T_\varepsilon > n) = \sigma_1(n) - \sigma_2(n) + \sigma_3(n) - \ldots + (-1)^n \sigma_{n+1}(n), \tag{2.11}$$

with

$$\begin{aligned} \sigma_k(n) = \; & \tbinom{n-1}{k-2}(1-(2k-2)\varepsilon)_+^n + 2\tbinom{n-1}{k-1}(1-(2k-1)\varepsilon)_+^n + \\ & + \tbinom{n-1}{k}(1-2k\varepsilon)_+^n, \end{aligned} \tag{2.12}$$

where $\binom{n}{k}$ represent the binomial coefficient "n choose k", with the condition that $\binom{n}{k}0$ except the case when $0 \leq k \leq n$ and $(\ldots)_+$ is zero when the quantity in brackets is negative. For example,

$$\begin{aligned} &\sigma_1(n) = 2(1-\varepsilon)^n + (n-1)(1-2\varepsilon)_+^n \\ &\sigma_2(n) = (1-2\varepsilon)_+^n + 2(n-1)(1-3\varepsilon)_+^n + \tfrac{(n-1)(n-2)}{2}(1-4\varepsilon)_+^n \\ &\ldots \\ &\sigma_n(n) = (n-1)(1-(2n-2)\varepsilon)_+^n + 2(1-(2n-1)\varepsilon)_+^n \\ &\sigma_{n+1}(n) = (1-2n\varepsilon)_+^n. \end{aligned}$$

The formula (2.11) is rather complex. Still, for $\varepsilon \in [\frac{1}{2}, 1]$ it becomes quite simple:

$$P(T_\varepsilon > n) = \begin{cases} 1, & \text{for } n = 0 \\ 2(1-\varepsilon)^n, & \text{for } n \geq 1. \end{cases}$$

For $\varepsilon \in \left(\frac{1}{3}, \frac{1}{2}\right]$ we get

$$P(T_\varepsilon > n) = \begin{cases} 1, & \text{for } n = 0, 1 \\ 2(1-\varepsilon)^n + (n-2)(1-2\varepsilon)^n, & \text{for } n \geq 1. \end{cases}$$

The smaller ε is, the more complex the formula becomes. For very small ε, even the Bonferoni inequalities cannot be of much help. Indeed, if estimating $P(T_\varepsilon > n)$ by

$$\sigma_1 - \sigma_2 \leq P(T_\varepsilon > n) \leq \sigma_1, \tag{2.13}$$

we get

$$\begin{aligned} 2(1-\varepsilon)^n + (n-2)(1-2\varepsilon)_+^n - 2(n-1)(1-3\varepsilon)_+^n - \tfrac{n(n-1)}{2}(1-4\varepsilon)_+^n \leq \\ \leq P(T_\varepsilon > n) \leq 2(1-\varepsilon)^n + (n-1)(1-2\varepsilon)_+^n. \end{aligned}$$

We can obtain an estimate in terms of known data. The random variables $U_{(1)}, \ldots, U_{(n)}$ split the segment $[0,1]$ into $n+1$ segments $[0, U_{(1)}], [U_{(1)}, U_{(2)}], \ldots, [U_{(n)}, 1]$. The maximal length of these segments is a random variable L_n whose distribution is known (see Wilkis [5, p. 238]). Obviously $\{L_n > 2\varepsilon\} \subseteq \{T_\varepsilon > n\}$. But

$$P(L_n > 2\varepsilon) = \mu(\{x \in S \mid x_1 \geq 2\varepsilon \text{ or } \ldots x_n \geq 2\varepsilon, x_1 + \ldots + x_n \leq 1 - 2\varepsilon\}).$$

Based on the preceeding results we infer

$$P(L_n > 2\varepsilon) = \binom{n+1}{1}(1-2\varepsilon)_+^n - \binom{n+1}{2}(1-4\varepsilon)_+^n + \ldots \\ \ldots + (-1)^n \binom{n+1}{n+1}(1-(2n+2)\varepsilon)_+^n, \tag{2.14}$$

and hence we have the estimate

$$\binom{n+1}{1}(1-2\varepsilon)_+^n - \binom{n+1}{2}(1-4\varepsilon)_+^n + \ldots + (-1)^n \binom{n+1}{n+1}(1-(2n+2)\varepsilon)_+^n \leq \\ \leq P(T_\varepsilon > n) \leq 2(1-\varepsilon)^n + (n-1)(1-2\varepsilon)_+^n.$$

In fact, the probability that the length L_n be greater than some value x is

$$P(L_n > x) = \binom{n+1}{1}(1-x)_+^n - \binom{n+1}{2}(1-2x)_+^n + \ldots \\ \ldots + (-1)^n \binom{n+1}{n+1}(1-(n+1)x)_+^n. \tag{2.15}$$

Since $\int_0^1 (1-kx)_+^n dx = \frac{1}{k(n+1)}$, it follows that the mean of L_n is

$$E(L_n) = \int_0^1 P(L_n > x)dx = \frac{1 + \frac{1}{2} + \ldots + \frac{1}{n+1}}{n+1}. \tag{2.16}$$

The Haussdorf distance $D_n := D([0,1], \{U_1, \ldots, U_n\})$ enters the calculation for obtaining an estimate for $P(T_\varepsilon > n)$ with very small ε, better than (2.13).

Let $G_n(x) := P(D_n > x) = P(T_x > n)$ (according to (1.9)). Taking into account the relations (2.10)-(2.12), we have

$$G_{n+1}(x) = \sum_{k\geq 0}(-1)^k \Big(\binom{n}{k-1}(1-2kx)_+^n + 2\binom{n}{k}(1-(2k+1)x)_+^n + \\ + \binom{n}{k+1}(1-2k+2)x)_+^n\Big). \tag{2.17}$$

Then

$$E(D_{n+1}) = \int_0^1 G_{n+1}(x)dx = \\ = \frac{1}{n+1}\sum_{k\geq 0}(-1)^k \left(\binom{n}{k-1}\frac{1}{2k} + 2\binom{n}{k}\frac{1}{2k+1} + \binom{n}{k+1}\frac{1}{2k+2}\right) = \\ = \frac{1}{n+1}\left[\frac{1}{2}\sum_{k\geq 0}(-1)^k \binom{n}{k-1}\frac{1}{k} + 2\sum_{k\geq 0}(-1)^k\binom{n}{k}\frac{1}{2k+1} + \frac{1}{2}\sum_{k\geq 0}(-1)^k\binom{n}{k+1}\frac{1}{k+1}\right].$$

We consider that $\binom{n}{k} = 0$ except of the case when $0 \leq k \leq n$; it follows that

$$\frac{1}{2}\sum_{k\geq 0}(-1)^k\binom{n}{k-1}\frac{1}{k} = -\frac{1}{2}\sum_{k\geq 0}(-1)^k\binom{n}{k+1}\frac{1}{k+1} = -\frac{1}{2(n+1)} \\ 2\sum_{k\geq 0}(-1)^k\binom{n}{k}\frac{1}{2k+1} = 2\frac{2\cdot 4 \cdot \ldots \cdot (2n)}{3 \cdot 5 \cdot \ldots \cdot (2n+1)} \\ \frac{1}{2}\sum_{k\geq 0}(-1)^k\binom{n}{k+1}\frac{1}{k+1} = \frac{1}{2}\left(1 + \frac{1}{2} + \ldots + \frac{1}{n}\right), \tag{2.18}$$

hence the mean of D_{n+1} is, for any $n \geq 1$, given by

$$E(D_{n+1}) = \tfrac{1}{n+1}\left[\tfrac{1}{2}\left(1+\tfrac{1}{2}+\ldots+\tfrac{1}{n}-\tfrac{1}{n+1}\right)+2\tfrac{2\cdot 4\cdot\ldots\cdot(2n)}{3\cdot 5\cdot\ldots\cdot(2n+1)}\right]. \tag{2.19}$$

We further obtain an estimate for $P(T_\varepsilon > n)$ with very small ε, better than in (2.13). For any natural $k \geq 1$, consider the net $G_k = \left\{\frac{j}{k} \mid 0 \leq j \leq k\right\}$. Let $T_\varepsilon^{(k)} = \max\limits_{0\leq j\leq k} T_{\varepsilon,\frac{j}{k}}$ the first moment when all the points of the net G_k have been bombed. Obviously, $T_\varepsilon \geq T_\varepsilon^{(k)}$.

Proposition 2.5. *Let $0 < \varepsilon \leq \frac{1}{2k}$. Then*

$$P(T_\varepsilon > n) \geq P(T_\varepsilon^{(k)} > n) \tag{2.20}$$

and

$$P(T_\varepsilon^{(k)} > n) = \theta_1 - \theta_2 + \ldots + (-1)^k\theta_{k+1}, \tag{2.21}$$

where

$$\begin{aligned}\theta_j = \;& \tbinom{k-1}{j-2}(1-2(2j-2)\varepsilon)^n + \\ & +2\tbinom{k-1}{j-1}(1-(2j-1)\varepsilon)^n + \tbinom{k-1}{j}(1-2j\varepsilon)^n.\end{aligned} \tag{2.22}$$

Proof. Consider the set

$$C_j = \left\{\frac{j}{k} \notin \bigcup_{m=1}^{n}[U_m-\varepsilon, U_m+\varepsilon]\right\} = \bigcap_{m=1}^{n}\left\{\left|U_- \frac{j}{k}\right| > \varepsilon\right\}.$$

Since the random variables U_m are independent and ε is sufficiently small, we notice that

$$P(C_j) = \begin{cases} (1-\varepsilon)^n, & \text{for } j \in \{0,k\} \\ (1-2\varepsilon)^n, & \text{for } 1 \leq j \leq k-1. \end{cases} \tag{2.23}$$

Taking into account that for $J \subseteq \{0,1,\ldots,k\}$ and $j = |J|$ we have

$$P\left(\bigcap_{\in J} H_i\right) = \begin{cases} (1-(2j-2)\varepsilon)^n, & \text{for } \{0,k\} \subset J \\ (1-(2j-1)\varepsilon)^n, & \text{for } |\{0,k\}\cap J| = 1 \\ (1-2j\varepsilon)^n, & \text{for } \{0,k\}\cap J = \varnothing\,, \end{cases} \tag{2.24}$$

it follows that (2.21) represents the Poincaré formula. $\square$

In the case when $\varepsilon = \frac{1}{t}$, where t is a positive integer, $t \geq 2$, we can determine an upper bound for the probability $P(T_\varepsilon > n)$. The idea relies on considering the net G_t an on waiting until all the intervals $I_j = \left[\frac{j}{t}, \frac{j+1}{t}\right)$ are bombed.

We define T'_ε as the first moment when the interval I is covered such that the segments I_j determined by the points of the net, segments of length greater than ε, contain a bomb center U_k, hence

$$T'_\varepsilon = \inf\{n \geq 1 \mid C_n \cap I_j \neq \varnothing\,, \forall 0 \leq j \leq t-1\}, \tag{2.25}$$

where the intervals I_j were defined above.

Proposition 2.6. *Assuming that the conditions stated above hold true, it follows that* $T_\varepsilon \le T'_\varepsilon$, *and hence*

$$P(T_\varepsilon > n) \le P(T'_\varepsilon > n). \tag{2.26}$$

Moreover,

$$P(T'_\varepsilon > n) = \binom{t}{1}(1-\varepsilon)^n - \binom{t}{2}(1-2\varepsilon)^n + \ldots + (-1)^{t-1}\binom{t}{t}(1-t\varepsilon)^n. \tag{2.27}$$

3 The mean of T'_ε: a calculation formula and an estimate

Proposition 3.1. Let $t = \frac{1}{\varepsilon}$. If $1 \le t \le 2$, then

$$E(T_\varepsilon) = 2t - 1. \tag{3.1}$$

If $t > 2$, then

$$E(T_\varepsilon) = \frac{t^2}{4} + t - \sum_{k \ge 2} (-1)^k \left(\left(\frac{t}{2k-2} - 1 \right)^{k-1} + 2\left(\frac{t}{2k-1} - 1 \right)^k + \left(\frac{t}{2k} - 1 \right)^{k+1} \right). \tag{3.2}$$

For example,

- if $2 \le t \le 3$, then $E(T_\varepsilon) = 1 + \frac{t}{2} + \frac{t^2}{4}$;
- if $3 \le t \le 4$, then $E(T_\varepsilon) = 1 + \frac{t}{2} + \frac{t^2}{4} - 2(\frac{t}{3} - 1)^2$;
- if $4 \le t \le 5$, then $E(T_\varepsilon) = 1 + \frac{t}{2} + \frac{t^2}{4} - 2(\frac{t}{3} - 1)^2 + (\frac{t}{4} - 1)^2 - (\frac{t}{4} - 1)^3$.

For the proof, we address to [1, pp. 50-51].

Remark 3.2. Let us denote $E(T_\varepsilon)$ by $E(t)$, where $t = \frac{1}{\varepsilon}$. Examining the expression of $E(t)$ given by (3.2), we remark that for $t \in [m, m+1)$, with m positive integer, $E(t)$ is a polynomial function. Its degree is given by the last k for which $t > 2k$. For t sufficiently big, the degree of $E(t)$ considerably increases.

For example, for $t = 100$, the degree of $E(t)$ is 50. Formula (3.2) has the disadvantage that the sum in the right side of the equality is hard to compute in the neighborhood of 0, since the alternating terms occuring in the sum are very big. Therefore, for a big value of t is necessary an estimate $E(t)$.

Proposition 3.3. *For any positive integer* $t > 2$, *the following inequalities hold true:*

$$\begin{aligned} \frac{t}{2}\log(\frac{t}{2} - 1) - \frac{2}{t-2} + \sqrt{2t} \quad &< E(t) < t\left(1 + \frac{1}{2} + \ldots + \frac{1}{t}\right) < \\ &< t(1 + \log t). \end{aligned} \tag{3.3}$$

Proof. According to Proposition 2.6. we have

$$P(T_\varepsilon > n) \le P(T'_\varepsilon > n) = \binom{t}{1}(1-\varepsilon)^n - \binom{t}{2}(1-2\varepsilon)^n + \ldots + (-1)^{t-1}\binom{t}{t}(1-t\varepsilon)^n.$$

Hence

$$
\begin{aligned}
E(t) \le E(T'_\varepsilon) &= \sum_{n\ge 0} P(T'_\varepsilon > n) = \\
&= -1 + \frac{\binom{t}{1}(1-\varepsilon)^n}{\varepsilon} - \frac{\binom{t}{2}(1-2\varepsilon)^n}{2\varepsilon} + \ldots + (-1)^{t-1}\frac{\binom{t}{t}(1-t\varepsilon)^n}{n\varepsilon} = \\
&= \left[1 - \tbinom{t}{1} + \tbinom{t}{2} - \tbinom{t}{3} + \ldots + (-1)^{t-1}\tbinom{t}{t}\right] + \\
&\quad + t\left[\tbinom{t}{1} - \tfrac{1}{2}\tbinom{t}{2} + \tfrac{1}{3}\tbinom{t}{3} + \ldots + (-1)^{t-1}\tfrac{1}{t}\tbinom{t}{t}\right] = \\
&= (1-1)^t + t\left(1 + \tfrac{1}{2} + \ldots + \tfrac{1}{t}\right) = t\left(1 + \tfrac{1}{2} + \ldots + \tfrac{1}{t}\right).
\end{aligned} \tag{3.4}
$$

On the other side, $T_\varepsilon \ge T_\varepsilon^{(k)}$, and hence $E(t) \ge E(T_\varepsilon^{(k)})$. We assume that $t \ge 2k$. Applying Proposition 2.5 and taking into consideration that

$$\sum_{n=0}^{\infty}(1 - m\varepsilon)^n = \frac{1}{m\varepsilon},$$

we get

$$E(T_\varepsilon^{(k)}) = \sum_{n=0}^{\infty} P(T_\varepsilon^{(k)} > n) = 1 + E_1 - E_2 + \ldots + (-1)^k E_{k+1}, \tag{3.5}$$

where

$$E_j = t\frac{\binom{k-1}{j-2}}{2(j-1)} + 2t\frac{\binom{k-1}{j-1}}{2j-1} + t\frac{\binom{k-1}{j}}{2j} - \binom{k}{j} \tag{3.6}$$

for any $j \ge 2$, with the condition that $\binom{k-1}{j} = 0$, except the case when $0 \le j \le k-1$. We can write (3.5) as $E(T_\varepsilon^{(k)}) = A + B + C + \sum_{j=0}^{k}(-1)^j\binom{k}{j} = A + B + C + (1-1)^k = A + B + C$, where

$$
\begin{aligned}
A &= -\tfrac{t}{2}\left[\tbinom{k-1}{0} - \tfrac{1}{2}\tbinom{k-1}{1} + \tfrac{1}{3}\tbinom{k-1}{2} - \ldots\right] = -\tfrac{t}{2}\tfrac{1}{k} \\
B &= 2t\left[\tbinom{k-1}{1} - \tfrac{1}{3}\tbinom{k-1}{2} + \tfrac{1}{5}\tbinom{k-1}{3} - \ldots\right] = 2t \cdot \tfrac{2\cdot 4\cdot\ldots\cdot(2k-2)}{3\cdot 5\cdot\ldots\cdot(2k-1)} \\
C &= \tfrac{t}{2}\left[\tbinom{k-1}{1} - \tfrac{1}{2}\tbinom{k-1}{2} + \tfrac{1}{3}\tbinom{k-1}{3} - \ldots\right] = \tfrac{t}{2}\left(1 + \tfrac{1}{2} + \ldots + \tfrac{1}{k-1}\right).
\end{aligned} \tag{3.7}
$$

It follows that

$$E(T_\varepsilon^{(k)}) = -\frac{t}{2k} + 2t \cdot \frac{2\cdot 4\cdot\ldots\cdot(2k-2)}{3\cdot 5\cdot\ldots\cdot(2k-1)} + \frac{t}{2}\left(1 + \frac{1}{2} + \ldots + \frac{1}{k-1}\right).$$

But since $\frac{2\cdot 4\cdot\ldots\cdot(2k-2)}{3\cdot 5\cdot\ldots\cdot(2k-1)} > \frac{\sqrt{2k+1}}{2k} > \frac{1}{\sqrt{2k}}$ for any $k \ge 2$ (the first inequality can be immediately verified by induction, and the second is obvious) and

$$\frac{t}{2}\left(1 + \frac{1}{2} + \ldots + \frac{1}{k-1}\right) > \frac{t}{2}\log k,$$

we infer

$$E(T_\varepsilon^{(k)}) > -\frac{t}{2k} + 2t\frac{1}{\sqrt{2k}} + \frac{t}{2}\log k$$

for any $t \geq 2k$. Replacing k by $\left[\frac{t}{2}\right]$, we obtain the inequality

$$E(T) \geq E(T_\varepsilon^{(k)}) > -\frac{t}{2}\left[\frac{t}{2}\right] + 2t\frac{1}{\sqrt{2\left[\frac{t}{2}\right]}} + \frac{t}{2}\log\left[\frac{t}{2}\right] > \frac{t}{2}\log\left(\frac{t}{2}-1\right) - \frac{2}{t-2} + \frac{2t}{\sqrt{2t}},$$

hence we have obtained exactly the estimate (3.3) for $E(t)$ □

Corollary 3.4. *The following inequalities hold true:*

$$\frac{1}{2} \leq \liminf_{t\to\infty} \frac{T_{1/t}}{t\log t} \leq \limsup_{t\to\infty} \frac{T_{1/t}}{t\log t} \leq 1.$$

Proof. The claim follows straightforward, passing to limit for $t \to \infty$ in the sequence of inequalities (3.3). □

Let us denote the lower and the upper bounds in (3.3) by m and M respectively. Then $m = \frac{t}{2}\log(\frac{t}{2}-1) - \frac{2}{t-2} + \sqrt{2t}$ and $M = t(1+\log t)$. As well, let us denote by $E'(t) = E(T'_\varepsilon)$ and by $E(t,[t/2]) = E(T_\varepsilon^{(k)})$ for $k = [t/2]$. Obviously, for any $t \geq 4$, the following inequalities hold true (according to the proof of the Proposition 2.3)

$$m \leq E(t,[t/2]) \leq E(t) \leq E'(t) \leq M.$$

The attached table contains all these data computed for 40 values of $t = \frac{1}{\varepsilon}$, included for big values of t (e.g., for $t = 500$, when the degree of $E(t)$ is 250 according to the Remark 3.2). Cook has computed the mean $E(t)$ ([3]) for 12 values of t up to $t = 200$, both for the unit interval and for the unit circle.

It can be seen, as expected, that the values increase on horizontal from left to right, and on vertical, from up down. The values of $E(t)$ seem to be closer to m than to M. Moreover, it can be noticed that, in the considered cases, where $E(t) < \frac{m+M}{2}$, but we cannot obviously draw any conclusion for the general case. As well, we can observe that the values of $E(t)$ are closer to $E(t,[t/2])$ than to $E'(t)$. In other words, the lower estimates are somewhat better than the upper ones.

In order to follow easier the evolution of the computed data for the 40 values of t, we include as well the graphics of m, $E(t,[t/2])$, $E(t)$, $E'(t)$, M with respect to the same system of axes, in which the abscyssa is the t-axis.

We notice that, indeed, the graphic of $E(t)$ is closer to the graphic of m than to the one of M.

References

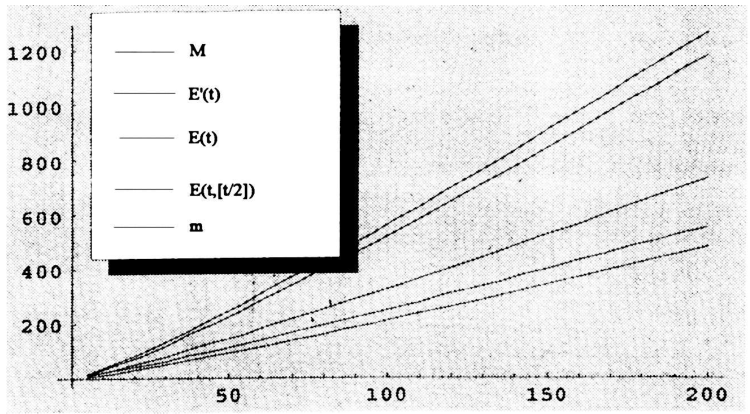
[1] Bontaş, S. (1998), *Studiul repartiţiilor timpilor de aşteptare într-un proces de acoperire a unui interval.* Teză de doctorat, Universitatea din Bucureşti.

[2] Bontaş, S., Zbăganu, Gh. (2000), *On the coupon collector problem in incompatible settings.* Rev. Roum. de Math. Pures et Appl., 55 (2000), 1, 49-66.

[3] Cooke, P.J. (1974), *Bounds for coverage probabilities with applications to sequel coverage problem.* J.Appl. Probability, 11, 281-293.

t	m	$E(t,[t/2])$	$E(t)$	$E'(t)$	M
1	x	x	1,000	1,000	1,000
2	x	x	3,000	3,000	3,386
3	-0,590	x	4,150	5,500	6,295
4	1,828	6,333	6,778	8,333	9,545
5	3,509	7,917	8,908	11,417	13,047
6	5,043	9,900	11,141	14,700	16,751
7	6,549	11,550	13,459	18,150	20,621
8	8,061	13,648	15,851	21,743	24,635
9	9,594	15,354	18,308	25,461	28,775
10	11,154	17,544	20,823	29,290	33,026
12	14,356	21,566	26,007	37,238	41,819
14	17,667	25,698	31,368	45,522	50,947
16	21,081	29,927	36,883	54,092	60,361
18	24,590	34,244	42,533	62,912	70,027
20	28,186	38,641	48,302	71,955	79,915
22	31,862	43,110	54,180	81,198	90,003
24	35,612	47,647	60,157	90,623	100,273
26	39,432	52,246	66,225	100,215	110,711
28	43,316	56,905	72,376	109,961	121,302
30	47,260	61,618	78,606	119,850	132,036
35	57,365	73,290	94,488	145,137	159,437
40	67,780	85,907	110,763	171,142	187,555
45	78,471	98,100	127,382	197,773	216,300
50	89,410	111,212	144,306	224,960	245,601
55	100,572	123,826	161,506	252,649	275,403
60	111,939	137,347	178,956	280,792	305,661
65	123,495	150,316	196,636	309,353	336,335
70	135,225	164,184	214,527	338,299	367,395
75	147,119	177,461	232,616	367,602	398,812
80	159,166	191,632	250,889	397,238	430,562
85	171,356	205,179	269,334	427,188	462,625
90	183,682	219,619	287,941	457,431	494,983
95	196,137	233,409	306,702	487,953	527,618
100	208,713	248,089	325,608	518,738	560,517
150	340,112	396,353	521,342	838,677	901,595
200	479,502	552,231	726,455	1175,610	1259,660
250	624,888	713,863	938,540	1525,170	1630,370
300	775,080	880,129	1156,180	1884,800	2011,130
400	1086,940	1223,770	1604,640	2627,970	2796,590
500	1410,980	1579,250	2067,010	3396,410	3607,300

The values of m, $E(t,[t/2])$, $E(t)$, $E'(t)$ and M in terms of t.



[4] Davy, P. (1982), *Coverage. Encyclopedia of Statistical Sciences.* Ed. S.Kotz and N.L. Johnson, vol. 2, 212-214, Wiley, N.Y.

[5] Wilks, S. (1962), *Mathematical Sciences.* Wiley, N.Y.

Silvia Bontaş
University Politehnica of Bucharest, Department Mathematics I
Splaiul Independenţei 313, RO-060042, Bucharest, Romania