Distance distributions in random networks

D. Moltchanov
Department of Communication Engineering,
Tampere University of Technology,
P.O.Box 553, Tampere, Finland
E-mail: moltchan@cs.tut.fi

Abstract

To account for stochastic properties when modeling connectivity in wireless mobile systems such as cellular, ad-hoc and sensor networks, spatial points processes are used. Since connectivity can be expressed as a function of distance between nodes distance distributions between points in spatial processes are of special importance. In this paper we survey those results available for distance distributions between points in two mostly used spatial point models, namely, homogeneous Poisson process in $\mathbb{R}^2$ and independently uniformly distributed points in a certain region of $\mathbb{R}^2$. These two models are known for decades and various distance-related results have been obtained. Unfortunately, due to wide application area of spatial point processes they are scattered among multiple field-specific journals and researchers are still wasting their time rediscovering them time after time. We attempt to unify these results providing an ultimate reference. We will also briefly discuss some of their applications.

Keywords: spatial point processes; distance distributions; connectivity

1 Introduction

Spatial point processes have been used in many branches of science to model various environmental phenomena. These areas include forestry, technology, geodesy, military applications and many more. Over more than one hundred years of extensive research and applications a lot of processes have been introduced and studied in detail. These processes also serve as a skeleton for some advanced structures in stochastic geometry, e.g. marked processes, coverage processes, tessellations, etc. For detailed introduction and review one could refer to [1, 2].
In the previous decade spatial points processes started to receive special attention from networking community in context of modeling spatial node distributions in various mobile wireless networks including ad-hoc, sensor and cellular networks. Random distribution of users on the plane allows to model special effects with connectivity being one of the most important. Since connectivity can be represented as a function of both distance between nodes and fading environment between them, the interest in networking community is rather limited and researchers are mostly concerned with various distance related metrics.

In this paper we concentrate on two simple models. The first model is homogeneous Poisson process on the plane. The second one is $N$ points independently and uniformly distributed (iud) in a certain region of the plane, including circle and rectangle. There are multiple reasons why these models are important in networking. First of all, in absence of extensive statistical studies of spatial nodes’ distribution in real networks the abovementioned models provide first-order approximation accounting for stochastic factors in the connectivity process. The second reason is more practical and related to tractability of these models. Given a large set of spatial point processes introduced in the past homogeneous Poisson process and iud points in a region allow for analytical characterization of distributions of various distance-related metrics.

The task of distance estimation between points of point processes in $\mathbb{R}^2$ is related to both theory of point processes and stochastic geometry. In both fields numerous studies have been carried to date. However, it happens that none of these fields is naturally concerned with distance distributions between points in $\mathbb{R}^2$. The reason is that the theory of point processes is mostly concerned with processes in one dimension for which a plenty of general results have been obtained. This interest is explained by rich application field of such results, e.g. reliability theory, traffic and teletraffic theories, etc. On the other hand, stochastic geometry is mostly concerned with models’ inference from statistical data for relatively complex spatial point models. Thus, the research on distributions of distances in point processes of dimensions higher than one have never been an issue of systematic research and have been performed in rather ad hoc way in the past.

In this paper we will try to organize those results available for distributions of distances between points for two most frequently used models in planar networking. For homogenous Poisson process in $\mathbb{R}^2$ we provide distribution of distance to the $n$th neighbor from an arbitrarily chosen point and joint distribution of distances to the first nearest $n$ neighbors. For better understanding whenever possible we will also provide analogy with one-dimensional homogenous Poisson process. For iud points in a region we provide expressions for distribution of the distance between two randomly chosen points, $n$th neighbor distance distribution from the origin of a circle, and distributions between nearest and farest points. We also provide results for the joint distance distribution between a chosen point and the
rest of \((N - 1)\) points in a region. When available extensions to \(\mathbb{R}^M\), \(M > 2\) are given. In other cases the complexity of extension to \(M > 2\) is discussed.

All abovementioned results have been rediscovered multiple times over the past century. One of the reasons is extensive application area of spatial models as these results have been published in many specific contexts. Now with extensive applications of spatial models in networking environment many authors still continue to derive them independently wasting their valuable time. The motivation behind this survey was to provide an universal reference for distance distribution for two mostly used spatial models in networking. This survey also contains elements of tutorial. The reason is that some available results are limited to a certain simple geometrical objects while underlying techniques can be extended to more complex structures. Ideas for using these results in applied networking studies are also given. Taking into account that some results have been known for decades we will also provide brief historical notes whenever appropriate.

The rest of the paper is organized as follows. In Section 2 we give basic characteristics of two spatial point models considered in this survey. Next, in Sections 3 and 4 distance distributions for these models are considered. Few applications are provided in Section 5. Finally, conclusions are given.

2 Spatial point processes

A process \(\Phi\) in \(\mathbb{R}^2\) is called homogenous (stationary) Poisson if it possess the following properties: (i) for any two disjoint Borel set \(A_1\) and \(A_2\) the random variables (RV) \(P(n, A_1)\) and \(P(n, A_2)\) describing the number of points of a process falling in these sets are independent (ii) the number of points \(P(n, A)\) falling in a bounded Borel set \(A\) is distributed according to Poisson law with parameter \(\lambda L(A)\), where \(L(A)\) is the area of \(A\), i.e.

\[
P(n, A) = \frac{[\lambda L(A)]^n}{n!} e^{-\lambda L(A)},
\]

where \(\lambda\) is the mean density (intensity) of points, \(E[P(n, A)] = \lambda L(A)\).

Note that the homogeneous point process is also isotropic, i.e. the process is the same with respect to rotation around an arbitrarily chosen origin. Homogeneity, isotropy and independence are three properties making analysis of Poisson process a relatively easy task. Notice that the latter property holds for non-homogenous Poisson processes too. Unless otherwise specified we assume these properties in what follows.

We also note that the property (i) is of special importance as there are no other spatial point processes with such strong independence property [2]. The second property implies that there are no multiple point occurring in the same
infinitesimally small area. Notice that this property is not required to be given as a part of definition. It is sufficient to assume that the probability that a point occurs in a region of area $dA$ is $\lambda dA$ and the probability that more than one point occurs in this area is of smaller magnitude than $\lambda dA$. Then, (1) immediately follows. To show it assume that $i$, $i \leq n$ points occur in $A$ and consider how we can have $n$ points in $A + dA$

$$P(n, A + dA) = P(n, A)P(0, dA) + \sum_{i=1}^{n} P(n - i, A)P(i, dA). \quad (2)$$

Observe that $P(0, dA) = 1 - \sum_{i=1}^{\infty} P(i, dA)$. Substituting it into (2), dividing by $dA$ and rearranging the terms we get

$$\frac{P(n, A + dA) - P(n, A)}{dA} = \frac{P(1, dA)}{dA} [P(n - 1, A) - P(n, A)] + \frac{P(2, dA)}{dA} [P(n - 2, A) - P(n - 1, A)] + \ldots. \quad (3)$$

Due to restriction on multiple occurrence of points $P(i, dA)/dA \to 0$ when $dA \to 0$. Noticing that $\lambda dA = P(1, dA)$ we have

$$\frac{P(n, A + dA) - P(n, A)}{dA} = \lambda [P(n - 1, A) - P(n, A)]. \quad (4)$$

Taking limits we have

$$\frac{dP(n, A)}{dA} + \lambda P(n, A) = \lambda P(n - 1, A). \quad (5)$$

Setting $n = 0$ in (5) we can see that the term in RHS disappears as $P(-1, A) = 0$ and we get get differential equation with boundary condition $P(0, 0) = 1$. Its solution is given by $P(0, A) = e^{-\lambda A}$. Setting $n = 1$ in (5) and substituting $P(0, A) = e^{-\lambda A}$ leads to the following differential equation

$$\frac{dP(1, A)}{dA} + \lambda P(1, A) = \lambda e^{-\lambda A}, \quad (6)$$

with boundary condition $P(1, 0) = 0$. The solution of (6) is $P(1, A) = \lambda A e^{-\lambda A}$. Continuing along these line we eventually approach Poisson distribution $P(n, A) = e^{-\lambda A}(\lambda A)^n/n!$ for $\lambda > 0$, $A > 0$, $n = 0, 1, \ldots$. The joint distribution of points in $m$ disjoint sets $A_i$, $i = 1, 2, \ldots, m$ ($m$-dimensional distribution) is given by

$$\left[\lambda L(A_1)\right]^{k_1}/k_1! \ldots \left[\lambda L(A_m)\right]^{k_m}/k_m! \exp \left(-\lambda \sum_{i=1}^{m} L(A_i)\right). \quad (7)$$
Let us now introduce another widely used model in ad-hoc networking. Let $W$ be a compact set, e.g. a circle (disk) or rectangle. Assume that $N$ points are iud in $W$. This implies that (i) all $N$ points are stochastically independent, i.e. the probabilities $Pr\{x_i \in A_i\}$, $A_i \subseteq W$, $i = 1, 2, \ldots, N$ satisfy the product expression

$$Pr\{x_1 \in A_1, \ldots, x_N \in A_N\} = Pr\{x_1 \in A_1\} \cdots Pr\{x_N \in A_N\}, \quad (8)$$

and (ii) each point is uniformly distributed in $W$, i.e. for any point $x_i$, $i = 1, 2, \ldots, N$ and $A \subseteq W$ we have

$$Pr\{x_i \in A\} = \frac{L(A)}{L(W)}, \quad (9)$$

that is, the probability that $x_i$ is in $A$ is proportional to the area of $A$.

The mean number of points per unit area is $\lambda = N/L(W)$, while the mean number of points in a set $A$ is $E[N(A)] = \lambda L(A)$. The one-dimensional distributions are of binomial form

$$P(k, N, A) = \binom{N}{k} p_A^k (1 - p_A)^{N-k}, \quad k = 0, 1, \ldots, N, \quad (10)$$

where $p_A = L(A)/L(W)$.

Finally, $m$-dimensional distributions are

$$\frac{N!}{k_1! \cdots k_m!} \frac{[L(A_1)]^{k_1} \cdots [L(A_m)]^{k_m}}{[L(W)]^N}, \quad (11)$$

where $A_i \cap A_j = 0$, for any $i$ and $j$, $i \neq j$, are pairwise disjoint sets such that $\bigcup_i A_i = W$. Notice that even for pairwise disjoint sets the number of points falling to them are not independent.

Now let us demonstrate an important relation between the homogenous Poisson process and the process of $N$ iud points in a region introduced above. Let us fix a certain area $W \subseteq R^2$ and assume that exactly $N$ points of a homogenous Poisson process fall in this area. Let us find the conditional distribution of this process. Assuming $A \subseteq W$ and denoting the number of points falling in $A$ by $N(A)$ for a certain $k \in \{0, 1, \ldots, N\}$ we have

$$Pr\{N(A) = k | N(W) = N\} = \frac{Pr\{N(A) = k, N(W) = N\}}{Pr\{N(W) = N\}} = \frac{Pr\{N(A) = k, N(W \setminus A) = N - k\}}{Pr\{N(W) = N\}}. \quad (12)$$
Using independence property the numerator of (12) is written as $Pr\{N(A) = k, N(W\setminus A) = N - k\} = Pr\{N(A) = k\} Pr\{N(W\setminus A) = N - k\}$. Substituting it into (12) and evaluating using (1) we get

$$Pr\{N(A) = k|N(W) = N\} = \frac{e^{-\lambda L(A)} (\lambda L(A))^k}{k!} \frac{e^{-\lambda L(W\setminus A)} (\lambda L(W\setminus A))^{N-k}}{(N-k)!} =$$

$$= \frac{N!}{k!(N-k)!} \left(\frac{L(A)}{L(W)}\right)^k \left(\frac{L(W\setminus A)}{L(W)}\right)^{N-k} =$$

$$= \binom{N}{k} p^k (1-p)^{N-k}, \quad (13)$$

where $p = L(A)/L(W)$.

Observing (13) we see that given $N(W) = N$ for any $A \subseteq W$ the number of points has binomial distribution. Recall, that a homogeneous spatial process has binomial distribution if the underlying point pattern is iud in $W$. In other words, given $N$ points of the Poisson process in $W$, these points are conditionally iud in $W$. Observe that the number of points in disjoint sets are no longer independent as occurrence of $k$ points in $A \subseteq W$ implies that there are exactly $(N - k)$ points in $W\setminus A$. This property makes analysis of distances much more complicated compared to the homogenous Poisson process.

This is backward connection between $N$ iud points in $W$ and the Poisson process in $\mathbb{R}^2$. When $N$ tends to infinity while $p_A$ tends to $0$ in (13) the distribution of the process is asymptotically Poisson. This limit is obtained when $W$ is enlarged to cover $\mathbb{R}^2$ while $N$ tends to infinity such that $N/L(W)$ is constant.

The conditional distribution property is used to generate Poisson process in simulation studies. Practically, in order to generate a Poisson process with intensity $\lambda$ in $W$ we first generate a Poisson variate with mean $\lambda L(W)$. For the resulting number, say $l$, we generate $l$ iud points in $W$. This binomial pattern in $W$ can be considered as a realization of a Poisson process in $W$. The difference is that in each attempt there will be different number of points of a Poisson process.

Another property of the Poisson point process is in special behavior of its Palm distributions. Palm distribution of a process is defined as distribution of points given that there is a point in the origin. For homogenous Poisson process this distribution coincides with original distribution of points (see e.g. [2] for more detailed discussion). Coupled with independence property this implies that various distance measures remain the same irrespective of whether they are estimated from an arbitrary point of the process or from an arbitrarily chosen location. Notice that this is not true for $N$ iud points in a region. For example, as we will see in what follows, the distance to the nearest point is different and depends on whether we condition it on the existence of a point at the origin. For further information on spatial processes we refer to [1], Ch. 13, [2] Ch. 2,4,5, and [3].
3 Poisson process in \( \mathbb{R}^2 \)

3.1 Void and contact distributions

Void (emptiness) and contact probabilities are the simplest distance-related characteristics of a point process. Void probability is defined as the probability of having no points in a certain test set \( A \), \( V(0, A) \). From the definition of the homogenous Poisson process it immediately follows that

\[
V(0, A) = P(0, A) = e^{-\lambda L(A)}.
\]  

(14)

Setting \( L(A) = \pi r^2 \) we get well-known results for circular void probabilities of a Poisson process. The void probabilities are tightly connected with another important distance-related quantity called contact distribution functions. If \( A \) is a Borel set with \( L(A) > 0 \) then the contact distribution function \( C(r, A) \) is

\[
C(r, A) = 1 - V(0, Ar) = 1 - e^{-\lambda L(A)}, \quad r \geq 0,
\]  

(15)

where \( rA \) is the dilation of \( A \) by a factor \( r \), i.e, \( rA = \{rx : x \in A\} \).

An important case is when \( A \) is the unit circle, \( A = \mathbb{A}_2(o, 1) \). Then,

\[
C(r, S) = 1 - e^{-\lambda \pi r^2},
\]  

(16)

is the so-called circular contact distribution. The corresponding density is

\[
c(r, S) = 2\lambda \pi r e^{-\lambda \pi r^2}, \quad r \geq 0.
\]  

(17)

As we will see (17) coincides with pdf of the distance from a random location \( o \) to the nearest point. Circular void probabilities and contact distribution are illustrated in Fig. 1. Notice that results similar to (14), (16) and (17) have straightforward interpretation in \( \mathbb{R}^1 \). Indeed, (14) gives probability that there are no point of a process in a segment of a certain length. (16) and (17) give CDF and pdf of the length of a segment till the next point on a line. Similar interpretations are available when time is considered as an index set of a process. Exponential nature of the contact distribution implies that the process is memoryless in nature. Due to this the structure of the distribution does not depend on whether \( o \) coincides with the point of a process or not.

3.2 Distance to \( n \)th neighbor

3.2.1 \( \mathbb{R}^2 \) space

Assume that we are given a Poisson process with intensity \( \lambda \). Let \( r_i, i = 1, 2, \ldots, n \) be the set of RVs denoting distances from a randomly chosen location up to the

\[\text{From now on, } o \text{ refers to the origin, while the index in } A_M \text{ to the dimension.}\]
nth point. We are interested in distributions of $r_i$, $i = 1, 2, \ldots, n$. Let $f(r, i)$ and $F(r, i)$, $r \geq 0, i = 1, 2, \ldots, n$ denote pdfs and CDFs of these distances.

The history of derivation of the solution is very rich. According to [4] it was Hertz who first solved this problem in 1909 for $n = 1$ (nearest neighbor problem) in arbitrary dimension [5]. Extensions for $n > 1$ have been provided by Chandrasekhar in 1943 [6], Skellam in 1951 [7] and Morishita in 1954 [8]. In [9] Thompson extended these results obtaining joint distribution of distances up to $n$th neighbor. It is interesting to note that at that time authors were mostly unaware of previous results in this field (see Thompson’s notes in [9]). Taking into account that most of those publications are not currently available on-line for detailed derivation and discussion we recommend [4]. Below, we basically follow [4] with some details of [9]. We also note that recently these results have been rediscovered multiple times.

Recall that the probability of having $n$ points in a region follows Poisson law. Let this region be a circle with area $\pi r^2$. Assume that the center of the circle $o$ is located at random and not necessarily coincides with a point of the process. Observe that if a circle of radius $r$ contains exactly $(n - 1)$ points and the $n$th point is located at the circumference, $r$ is the distance to the $n$th neighbor of $o$. The probability that a circle contains exactly $n$ points is

$$P(n, A) = \frac{\lambda L(A)^n e^{-\lambda L(A)}}{n!}.$$

The probability of finding at least $n$ points in a circle is given by

$$P(\geq n, A) = 1 - \sum_{i=0}^{n-1} P(n, A) = 1 - \sum_{i=0}^{n-1} \frac{[\lambda L(A)]^i}{i!} e^{-\lambda L(A)}, \quad (18)$$

where $L(A) = \pi r^2$.

The probability that the $n$th nearest point is found in the interval $(r + \Delta r)$ equals to the probability that this point is located in the annulus with inner radius $r$ and outer radius $(r + \Delta r)$. Thus,

$$P(n, n \in \pi(r + \Delta r)^2 - \pi r^2) = P(n, \pi(r + \Delta r)^2) - P(n, \pi r^2), \quad n \geq j. \quad (19)$$
Letting $\Delta r \to 0$ we get
\[ P(n, n \in \pi (r + \Delta r)^2 - \pi r^2) = P(n, \pi (r + \Delta r)^2) - P(n, \pi r^2). \]  
(20)

Differentiating with respect to $r$ we get
\[ f(r, n) dr = \frac{2(\pi \lambda)^n}{(n-1)!} r^{2n-1} e^{-\pi \lambda r^2} dr, \quad r > 0, \ n = 1, 2, \ldots. \]  
(21)

Notice that (21) gives pdf of the distance from an arbitrarily chosen origin $o$ to the $n$th nearest point. Recall, that due to memoryless property of Poisson process this gives the distance from an arbitrary point of the process to its $n$th neighbor, i.e. the centroid $o$ may or may not coincide with a point.

There is straightforward analogy with one-dimensional Poisson process. Recall that for homogenous Poisson process in $\mathbb{R}^1$ the distance or time duration from an arbitrary location or time instant till $n$th event is known to have Erlang distribution of order $n$. Observe now the structure of (21). Setting $t = \pi \lambda r^2$ and substituting in (21) gives the classic form of gamma distribution with integer valued shape parameter $(n-1)!$, i.e.
\[ f(t, n) = \frac{t^{n-1}}{(n-1)!} e^{-t}, \quad t > 0, \ n = 1, 2, \ldots. \]  
(22)

More precisely, (22) is Erlang distribution (of order $n$) which is widely used in teletraffic theory as the time interval up to $n$th arrival from a Poisson call arrival process. Furthermore, setting $t = 2\pi \lambda r^2$ and substituting into (21) results in classic form of chi-square distribution with $2n$ degrees of freedom. Moments and CDF for both chi-square and Erlang distributions are readily available elsewhere and for this reason not given here.

### 3.2.2 Extension to $\mathbb{R}^M$

In 2005 Haenggi [10] extended (21) to $\mathbb{R}^M$. We will briefly highlight his results here. It is important to stress before we proceed that the stricture of the distribution remains unchanged across all $M$. The only difference is coefficients of (21) that depends on the dimension we are working with.

Let $A_M(r) = c_M r^M$ be the $M$-dimensional spherical region$^2$ of radius $r$, where $c_M$ is given by
\[
c_M = \begin{cases} 
\frac{\pi^{M/2}}{(M/2)!}, & M = 2, 4, \ldots, \\
\frac{\pi^{M-1}}{2^M (M-1)!} \frac{M!}{M!}, & M = 1, 3, \ldots,
\end{cases}
\]  
(23)

$^2$Following [10] we will call it a $M$-dimensional ball.
e.g. for $M = 1$ and $M = 2$ we have $c_1 = 2$ and $c_2 = \pi$, respectively.

Let further $P(k, A_M(r)) = [\lambda A_M(r)]^k/k!$, $k = 0, 1, \ldots$ be the elements of Poisson distribution and let $r_n$ be the distance to the $n$th nearest point. Similarly to $R^2$ the probability that there are less than $n$ nodes closer than $r$ is

$$P(\geq n, A_M(r)) = 1 - \sum_{k=0}^{n-1} P(k, A_M(r)) e^{-\lambda A_M(r)}. \quad (24)$$

Further Haenggi proceeds using CDFs, defined as $P(< n, A_M(r)) = 1 - P(\geq n, A_M(r))$. Similarly to $R^2$ we have

$$f(r, n) = \lambda c_M M r^{M-1} \left( \sum_{k=0}^{n-1} P(k, A_M(r)) - \sum_{k=1}^{n-1} P(k-1, A_M(r)) \right) e^{-\lambda A_M(r)} =$$

$$= \frac{nM}{r} P(n-1, A_M(r)) e^{-\lambda A_M(r)} = \frac{nM}{r} P(n, A_M(r)) e^{-\lambda A_M(r)}. \quad (25)$$

Substituting Poisson probabilities and rearranging we get

$$f(r, n) = \frac{M(\lambda c_M)^n}{r \Gamma(n)} e^{-\lambda c_M r^M}, \quad r > 0, \ n = 1, 2, \ldots. \quad (26)$$

which is again pdf of generalized Gamma distribution with integer valued shape parameters whose CDF and moments are well-known.

Haenggi also mentions one straightforward property of homogenous Poisson process with intensity $\lambda$ in $R^M$. Assuming that points $X_i, i = 1, 2, \ldots$ of a Poisson process are ordered according to their Euclidian distance the distance $r_i = ||y-X_i||$ has the same distribution as in one-dimensional Poisson process with intensity $\lambda c_M$. For $M = 2$ this result has been mentioned in e.g. [11] among others. Additionally, $r_1$ as well as $(r_i - r_{i-1})$, $i > 1$ are similarly and exponentially distributed with mean $1/\lambda c_M$ and $E[r_i^M] = i/\lambda c_M$.

### 3.3 Joint distance distribution to $n$th neighbors

To the best of our knowledge Thomson was the first who studied joint distance distribution up to the $n$th neighbor [9]. From our previous discussion it is easy to see that the probability that there are no points in a circle of radius $r_1$ is $P(0, r_1) = e^{-\lambda \pi r_1^2}$ (see circular void probability). The probability that at least one point occurs in the annulus of circles with radii $r_1$ and $(r_1 + \Delta r_1)$ is $(1 - e^{-\lambda (\pi \Delta r_1)})$ and converges to $2\pi \lambda r_1 \Delta r_1$ as $\Delta r_1 \to 0$. Thus, the pdf of the distance to the nearest neighbor is $f(r, 1) = e^{-\lambda \pi r_1^2} 2\lambda r_1 dr_1$. We will denote it by $f(r_1)$ in this
Subsection. Although this result immediately follows from (21) setting \( n = 1 \) we will need this discussion further.

Let us now define \( f(r_2|r_1) \) to be the probability that the second nearest neighbor is at the distance \( r_2 \) given that the nearest one is at the distance \( r_1 \). Let further \( f(r_1, r_2) \) be the joint pdf of distances to the two nearest points. Similarly to the previous discussion for probability of having no points in between \( r_1 \) and \( r_2 \) and at least one point (or equivalently exactly one for Poisson process) in \( (r_2 + \Delta r_2) \) we can write

\[
P(0, A \in (\pi r_2^2 - \pi r_1^2)) = e^{-\lambda \pi (r_2^2 - r_1^2)},
\]

\[
P(> 0, A \in (\pi (r_2 + \Delta r_2) - \pi r_2^2)) = 1 - e^{-\lambda \pi \Delta r_2} = 2\lambda \pi r_2 dr_2.
\]

Using these results for conditional and joint pdfs we have

\[
f(r_2|r_1) = e^{-\lambda \pi (r_2^2 - r_1^2)}2\lambda \pi r_2 dr_2,
\]

\[
f(r_1, r_2) = f(r_1) f(r_2|r_1) = e^{-\lambda \pi r_2^2}(2\lambda \pi)^2 r_1 r_2 dr_1 dr_2.
\]

Using similar arguments for \( f(r_3|r_1, r_2) \) and \( f(r_1, r_2, r_3) \) we have

\[
f(r_3|r_1, r_2) = e^{-\lambda \pi (r_3^2 - r_2^2)}2\lambda \pi r_3 dr_3,
\]

\[
f(r_1, r_2, r_3) = e^{-\lambda \pi r_3^2}(2\lambda \pi)^3 r_1 r_2 r_3 dr_1 dr_2 dr_3.
\]

Continuing along these lines we approach the following joint pdf

\[
f(r_1, r_2, \ldots, r_n) = e^{-\lambda \pi r_n^2}(2\lambda \pi)^n r_1 \ldots r_n dr_1 \ldots dr_n.
\]

Notice that (30) gives joint distribution of distances. Obviously, the usage of (30) is rather tedious when \( n \) is large. Finally, note that (21) can be obtained from (30) by integrating it successively with respect to \( r_1 \) from \( o \) to \( r_2 \), \( r_2 \) from \( o \) to \( r_3 \), up to \( r_{n-1} \) from \( o \) to \( r_n \).

4 Uniform distribution in a region

4.1 Preliminaries

There are two ways to determine distances between iid points in a region \( A \). These are exact and approximate methodologies. Although the exact approach may sometimes be complicated as we will see in what follows (tedious algebraically even for two dimensions considered here) in this paper we do not consider approximate analysis. Some insights on approximate methodology for determining
distances between arbitrary points in a rectangle are given in [12]. The most important advantage of analytical techniques is that results are exact and there is no need for their further statistical inspection. Once obtained such results can be reused in many applications.

Before we proceed with special problems of interest let us consider the general methodology. Let \(D(x_1, y_1)\) and \(D(x_2, y_2)\) be two points of interest randomly distributed in \(A \subset \mathbb{R}^2\). Assume that we are interested in the distribution of Euclidean distance between these points, i.e. \(||D(x_1, y_1) - D(x_2, y_2)||\). Introduce the following distributions:

- spatial node distribution with pdf \(f(x, y)\) and CDF \(F(x, y)\);
- difference distribution, with pdf \(f_\Delta(\Delta X, \Delta Y)\) and CDF \(F_\Delta(\Delta X, \Delta Y)\);
- Euclidian distance distribution, with pdf \(f_L(l)\) and CDF \(F_L(l)\).

Functions \(f(x, y)\) and \(F(x, y)\) describe joint density and distribution of \(x\) and \(y\) coordinates of a point in \(A\). The difference pdf \(f_\Delta(\Delta X, \Delta Y)\) and CDF \(F_\Delta(\Delta X, \Delta Y)\), are the pdf and CDF of the difference between respective coordinates of two points, \(\Delta X = ||x_1 - x_2||\), \(\Delta Y = ||y_1 - y_2||\). The Euclidian distance distribution describes the distance between two points. In Cartesian coordinates it is given by \(||L||_C = \sqrt{\Delta X^2 + \Delta Y^2}\). Notice that working with polar coordinates is not generally easier even in special cases e.g. when the area of interest is circle. The reason is that the distance between two points is expressed as \(||L||_P = \sqrt{x_1^2 + x_2^2 - 2x_1x_2\cos(\theta_1 - \theta_2)}\). Thus, in order to determine the distance between two points instead of difference distribution we need to determine the distribution \(x_1x_2\cos(\theta_1 - \theta_2)\) and then the distribution of \(||L||_P\). The latter involves dependent components even when \(x\) and \(y\) coordinates are i.i.d and is more complicated to estimate compared to \(||L||_C\). However, using well-known relations \(x = r\cos \theta_1, y = r\sin \theta_1\) polar coordinates can be always translated to Cartesian ones. This discussion is also true for other coordinate systems, e.g. cylindrical.

Given the abovementioned distributions the analytical approach is as follows. Firstly, one determines the distribution of the coordinate difference \(\Delta X\) and \(\Delta Y\) using algebraic methods. Then, using well-known expressions for a power function of a RV one needs to find distributions \(\Delta X^2\) and \(\Delta Y^2\). At the next step integration is used to obtain the distribution of difference \((\Delta X^2 - \Delta Y^2)\). At the final step one uses expressions for distribution of the square root function of a RV to determine Euclidian distance between points. The most difficult part of the procedure is estimation of the double integral over circular region that is needed to find the distribution of \((\Delta X^2 - \Delta Y^2)\).

In some specific cases and, especially, when mean values of distances are of interest the easier solution is provided by the Crofton mean value and fixed points.
theorems. In what follows, we formulate them without proofs in the form they appeared in [11, 13]. These theorems show a way how to evaluate definite integrals without performing direct integration. A number of complicated problems can be easily solved applying these two theorems.

- Fixed points theorem, (Crofton, 1885, [14]).
  Let \( n \) points \( \epsilon_i, i = 1, 2, \ldots, n \), be randomly distributed in \( A \) and let \( H \) be some event (property) that depends on the positions of these points. Let \( A' \subset A \) such that \( \delta A \) is a part of \( A \) not in \( A' \). Then the following relation can be used to find the probability of certain point arrangements
  \[
dPr\{H\} = n(P\{H|\epsilon_1 \in \delta A\} - P\{H\})A^{-1}dA.
\] (31)

- Mean value theorem (Crofton, 1885, [14])
  Let \( X \) be a RV that depends on the position of \( n \) random points \( \epsilon_i, i = 1, 2, \ldots, n \), in \( A \). Using assumptions of the previous theorem we have the following relation for the increment of \( E[X] \)
  \[
dE[X] = n(E[X|\epsilon_1 \in \delta A] - E[X])A^{-1}dA.
\] (32)

We will illustrate the usage of the fixed points theorem. An interested reader is referred to [15], [11] Ch.2 and [13], Ch.5 for proofs, further reading and additional examples including widely known Sylverster’s four points problem.

### 4.2 Void and contact probabilities

Let \( \Phi \) be the stationary Poisson process in \( \mathbb{R}^2 \). Considering a compact set \( W \) let a new process to be defined as \( \Phi(W) = N \), a process having exactly \( N \) points in \( W \). As we already know these \( N \) points are iid in \( W \). Let also \( A \) be a compact subset of \( W \), \( A \subseteq W \). Then, the void probability of the conditioned stationary Poisson process is

\[
Pr\{\Phi(A) = 0|\Phi(W) = N\} = \frac{Pr\{\Phi(A) = 0, \Phi(W) = N\}}{Pr\{\Phi(W) = N\}} = \frac{Pr\{\Phi(A) = 0, \Phi(W\setminus A) = N\}}{Pr\{\Phi(W) = N\}} = \frac{[L(W) - L(A)]^N}{[L(W)]^N},
\] (33)

which coincides with void probabilities of the binomial point process. According to the general theory of point processes if void probabilities of two point processes
coincide then their distributions coincide too (see e.g. [2]). Thus, (33) shows that \( \Phi(W) = N \) is a special case of the binomial point process.

An important case is when \( A \) is the circle of radius \( r \), \( A = A_2(o, r) \). We have

\[
V(0, A) = \left(1 - \left(\frac{r}{R}\right)^2\right)^N,
\]

which is called circular void probability for \( N \) i.i.d points in \( W \).

Result (34) can be easily extended to \( M \)-dimensional space. Assuming that the test set \( A \) is \( M \)-dimensional ball of radius \( r \), i.e. \( A = A_M(o, r) \), the void probability is given by [16]

\[
V(0, A) = \left[1 - \left(\frac{r}{R}\right)^M\right]^N.
\]

Contact distributions and, particularly, circular contact distribution, are more complicated to obtain compared to the homogenous Poisson process in \( \mathbb{R}^2 \). Recall, that the contact distribution gives the distribution of the distance from a certain origin \( o \) to the first neighbor. As we will see in what follows it depends on the form of an area where points are distributed and on the choice of the origin. We will show how to obtain it for both circle and rectangle as a part of the sets of distance distributions to the \( n \)th neighbor for these shapes.

4.3 Two random points

4.3.1 Circle

Consider now a circle \( A \) of radius \( r \). Assume that two points are randomly thrown in \( A \) and we are interested in pdf of the Euclidian distance between them, \( f_L(l, r) \). Note that the result for \( f_L(l, r) \) has been known for years. Unfortunately, we were not able to find the reference of its first appearance. The earliest mention we have found is [17], where the authors stated that "this result must be well-known but we have not been able to trace a reference to it". According to Blumenfeld in [18], Ch.10 the earliest references are due to Garwood [19] and Garwood and Tanner [20]. However, as we will see in what follows it is reasonable to assume that this result has been known to Crofton in 1885.

To solve this problem we use Crofton fixed points theorem. Let \( P \) denote the probability that two points are separated by a distance between \( l \) and \( l + \Delta l \) (see Fig. 2(a)). \( P_1 \) denotes the same probability given that one of the points is on the circumference of the circle. Thus, in our case (31) simplifies to

\[
dP = 2(P_1 - P)A^{-1}dA,
\]

(36)
where \( A \) is the area of the circle, i.e. \( A = \pi r^2 \) and \( dA = 2\pi rdr \).

Observe Fig. 2(a), where illustration of the problem is presented. When one point is on the circumference \( dA \), for two points to be separated by \( l \) another point must be exactly \( l \) distance away. This implies it should reside on a section of an annulus. When \( dl \) is infinitesimally small the area of the annulus is \( 2\phi ldl \), where \( \phi \) is readily found to be \( \cos^{-1}(l/2r) \). Thus, \( P_1 \) can be found as

\[
P_1 = \frac{2ldl \cos^{-1}(l/2r)}{\pi r^2}.
\] (37)

Figure 2: Illustration of the distance between random points.

Substituting (37) into (36) we get

\[
dP = 2 \left( \frac{2ldl \cos^{-1}(l/2r)}{\pi r^2} \right) \frac{2dr}{r}.
\] (38)

Rearranging terms gives

\[
rdP + 4Pdr = \frac{8ldl \cos^{-1}(l/2r)}{\pi r^2}.
\] (39)

Multiplying both sides by \( R^3 \) leads to

\[
r^4dP + 4r^3Pdr = \frac{r8ldl \cos^{-1}(l/2r)}{\pi}.
\] (40)

Integrating both sides we get

\[
Pr^4 = \frac{2l^2dl}{\pi} \int \frac{2r}{l} \cos^{-1} \left( \frac{l}{2r} \right) dr =
\]

\[
= \frac{2l^2dl}{\pi} \left( 4r^2 \cos^{-1} \left( \frac{l}{2r} \right) - 2lr \sqrt{1 - \frac{l^2}{4r^2}} \right) + C.
\] (41)
where $C$ is the integration constant that needs to be determined.

For $r = l/2$ two points have to fall on the circumference diametrically across. This event has probability 0. Therefore, for $r = l/2 P = 0$. Substituting this into (41) we get $C = 0$. Finally, pdf of the distance is given by

$$f_L(l, r) = \frac{2l}{r^2} \left( \frac{2}{\pi} \cos^{-1} \left( \frac{l}{2r} \right) - \frac{l}{\pi r} \sqrt{1 - \frac{l^2}{4r^2}} \right), \quad 0 < l < 2r. \quad (42)$$

Integrating it we get CDF

$$F_L(l, r) = 1 + \frac{2}{\pi} \left( \frac{l^2}{r^2} - 1 \right) \cos^{-1} \left( \frac{l}{2r} \right) - \frac{l}{\pi r} \left( 1 + \frac{l^2}{2r^2} \right) \sqrt{1 - \frac{l^2}{4r^2}}. \quad (43)$$

Mean and variance can be directly computed from (42) or (43)

$$E[L] = \frac{128r}{45\pi}, \quad \sigma^2[L] = r^2 - \left( \frac{128r}{45\pi} \right)^2. \quad (44)$$

Notice that the usage of Crofton theorem does not always result in easier computation. The complexity depends on the type of a figure we need to deal with. Particularly, it depends whether $Pr\{H|\epsilon_1 \in \delta A\}$ in (31) results in a simple expression. This is indeed true for a circle considered above. It is also the case when $M$-dimensional balls are considered. When a cube in $R^2$, $M > 1$ is considered $Pr\{H|\epsilon_1 \in \delta A\}$ depends on where a point falls on the side. Recall that due to complete symmetry we did not have this problem working with a circle. Practically, in order to find $Pr\{H|\epsilon_1 \in \delta A\}$ we need to consider a number of special cases as will be clear in what follows. While we could handle $M = 2$ the complexity of expression for $Pr\{H|\epsilon_1 \in \delta A\}$ grows with $M$. When a rectangle is of interest we also need to take into account which side of it a point falls on. In this case even for $M = 2$ expression for $Pr\{H|\epsilon_1 \in \delta A\}$ is hard to obtain. In both abovementioned cases straightforward integration is easier to perform.

Finally, we would like to notice that the Crofton theorem on fixed points has been extended in 1973 by Ruben and Reed [21], who considered the case where points are chosen at random in each of a number of domains (see [13] for details). An interesting note on the theorem highlighting equivalency of the Crofton theorem and conditioning techniques for computation of geometrical probabilities has been recently published [22]. Note that [22] also inspects deficiencies of the Crofton’s formulation of the theorem and provides additional examples.

### 4.3.2 Square and rectangle

Consider two random points $D(x_1, y_1)$ and $D(x_2, y_2)$ thrown on a rectangle with sides $B$ and $C$. We are interested in the distribution of the Euclidian distance between these points. In what follows, we use direct analytical approach to estimate
it. Although for simplicity we will assume $B = C$, derivation for $B \neq C$ is similar and is only slightly more tedious. Here, we basically follow the recent work of Kostin [23]. However, this result has been known for years, see e.g. [24].

Firstly, concentrate on the coordinate difference $\Delta X = (x_1 - x_2)$ and let $f(\Delta X)$ be its pdf. Assuming that the left bottom corner of a rectangle is located at the center of coordinates as shown in Fig. 2(b), $x_1$ and $x_2$ are iud over $[0, B]$. Since $(x_1 - x_2)$ can be represented as $[x_1 + (-x_2)]$ and recalling that the sum of independent RVs is obtained using convolution, we have

$$ f(\Delta X) = \int_{-\infty}^{\infty} f(x) f(x - \Delta X) dx, \quad (45) $$

where $f(x) = 1/B$ is the common pdf for both $x_1$ and $x_2$.

Observe that $f(x)$ is not zero in $[0, B]$, while $f(x - \Delta X)$ is not zero for $0 \leq (x - \Delta X) \leq B$. Thus, (45) needs to be solved for $(-B, \Delta X]$ when $\Delta X \leq 0$ and $(\Delta X, B]$ for $\Delta X > 0$. We have

$$ f(\Delta X) = \begin{cases} 0, & \Delta X \notin (-B, B], \\
\frac{\Delta X + B}{B^2}, & \Delta X \in (-B, 0], \\
\frac{B - \Delta X}{B^2}, & \Delta X \in (0, B]. \end{cases} \quad (46) $$

Recalling the expression for Euclidian distance we further need pdf of $\Delta X^2$. Assume we are given a RV $X$ with pdf $f(x)$ and let $Y = \phi(X)$ be another RV we are interested in with pdf $g(y)$ and CDF $G(y)$. If $\phi(x)$ is monotonous and differentiable in $[a, b]$ pdf and CDF of $Y$ are given by (see e.g. [25] more details)

$$ g(y) = \begin{cases} f(\phi(y)) \frac{d\phi(y)}{dy}, & \text{if } \frac{d\phi(x)}{dx} > 0 \forall x \in [a, b] \\
-f(\phi(y)) \frac{d\phi(y)}{dy}, & \text{if } \frac{d\phi(x)}{dx} < 0 \forall x \in [a, b]. \end{cases} \quad (47) $$

$$ G(y) = \begin{cases} \int_{a}^{\phi(y)} f(x) dx, & \text{if } \frac{d\phi(x)}{dx} > 0 \forall x \in [a, b] \\
\int_{\phi(y)}^{b} f(x) dx, & \text{if } \frac{d\phi(x)}{dx} < 0 \forall x \in [a, b]. \end{cases} \quad (48) $$

where conditions imply that a function is either increasing or decreasing in $[a, b]$.

Observe that (46) is monotonously increasing in $\Delta X \in (-B, 0]$ and monotonously decreasing in $\Delta X \in (0, B]$ (in fact, (46) is Simpson triangular distribution, [25]). Obviously, $Pr\{\Delta X^2 \leq 0\} = 0$ and $Pr\{\Delta X \geq B^2\} = 0$. Thus, pdf of $\Delta X^2$ is

$$ f(\Delta X^2) = \begin{cases} 0, & \Delta X^2 \notin (0, B^2], \\
\frac{1}{B\sqrt{\Delta X}} - \frac{1}{B^2}, & \Delta X^2 \in (0, B^2]. \end{cases} \quad (49) $$
Now we are ready to determine the distribution of $(\Delta_X^2 + \Delta_Y^2)$. First, observe
that RVs $\Delta_X^2$ and $\Delta_Y^2$ are independent of each other. The distribution of $\Delta_Y^2$ is
(46). Thus, denoting $u = (\Delta_X^2 + \Delta_Y^2)$ the CDF is formally written as

$$F_U(u) = \int_A f(\Delta_X^2) f(\Delta_Y^2) d\Delta_X^2 d\Delta_Y^2,$$  

(50)

where $A$ is the area below $u = (\Delta_X^2 + \Delta_Y^2)$ inside the square of size $B^2 \times B^2$.

Solving (50) we get

$$F_U(u) = \begin{cases} 
\pi a - \frac{8}{3} a^{3/2} + \frac{a^2}{2}, & u \in [0, B^2), \\
1 - \left(\frac{2}{3} + 2a + \frac{a^2}{2} - \frac{2(a-1)^{3/2}}{3}\right) - 2\sqrt{a-1} - 2a \arcsin \frac{2-a}{a}, & u \in [B^2, 2B^2),
\end{cases}$$  

(51)

where $a = u/B^2$. Note that $F_U(u) = 0$, $u < 0$, and $F_U(u) = 1$, $u \geq 2B^2$.

Finally, we need to determine $F_L(l)$, $l = \sqrt{u} = \sqrt{\Delta_X^2 + \Delta_Y^2}$. Observe that

$$F_L(l) = Pr\{L < l\} = Pr\{0 < u < l^2\}, \quad l \in (0, L\sqrt{2}).$$  

(52)

Noticing that $Pr\{0 < u < l^2\} = \int_0^{l^2} f(u) du$ and using (47) we get

$$F_L(l) = \begin{cases} 
\pi l^2 - \frac{8l^3}{3B^3} + \frac{l^4}{2B^4}, & l \in [0, B), \\
1 - \left(\frac{2}{3} + 2b + \frac{b^2}{2} - \frac{2(b-1)^{3/2}}{3}\right) - 2\sqrt{b-1} - 2b \arcsin \frac{2-b}{b}, & l \in [B, B\sqrt{2}),
\end{cases}$$  

(53)

where $b = l^2/B^2$. Note that $F_L(l) = 0$, $l < 0$, and $F_L(l) = 1$, $l \geq B\sqrt{2}$.

4.3.3 Higher dimensions and other shapes

Distribution of the distance between two random points in more than two dimensions has been recently addressed. Observe that the distance distribution between two points in a $M$-dimensional ball can be readily obtained applying the Crofton’s fixed points theorem. Mathai et al. derived distribution of Euclidian distance in a cube [26]. Similar result has been obtained by Philip in [27], who also extended it to 4th and 5th dimensions in [28]. Notice that extensions to even higher dimensions seems feasible. The limiting factor is increase of algebraic complexity associated with derivation procedure as clearly seen from the studies of Philip.
Non-Euclidian metrics including Manhattan and Chebychev ones have been considered in [29], where authors concentrated on mean values in rectangles and parallelepipeds. The review above is by no means exhaustive. However, an interested reader will find additional references in those papers.

Finally, we would like to note that the distances between random points distributed in other planar objects have not been deeply studied so far. The major reason is that for more complex shapes deriving expressions is tedious algebraically. However, this is still possible. An interested reader may refer to [30] for further discussion about the complexity of distance estimation in more complex objects and for expression for the mean distance between two points iud in a polygon.

4.4 Minimum and maximum distance

Assume now that there are \( N \) points iud in a rectangle or circle. We are interested in the minimum and maximum distance between these points. Since we choose a point randomly and the rest of points are iud, CDFs of the distance between a chosen point and other points are given by (43) and (53) for square and circle, respectively. Denote these distances by \( L_i, i = 1, 2, \ldots, N - 1 \). Let \( L_{\min} \) and \( L_{\max} \) be the distance to the nearest neighbor and to the most distant point and let \( F_{L_{\min}}(l) \) and \( F_{L_{\max}}(l) \) be their CDFs. Due to independence property we have

\[
L_{\min} = \min(L_1, L_2, \ldots, L_{N-1}), \\
L_{\max} = \max(L_1, L_2, \ldots, L_{N-1}).
\] (54)

The CDF of the minimum and maximum of a number of iid RVs is given by

\[
F_{L_{\min}} = 1 - [1 - F_L(l)]^{N-1}, \\
F_{L_{\max}} = [F_L(l)]^{N-1}.
\] (55)

Corresponding pdfs are obtained by differentiating

\[
f_{L_{\min}} = (N - 1)[1 - [F_L(l)]^{N-2}f_L(l), \\
f_{L_{\max}} = (N - 1)f_L(l)[F_L(l)]^{N-2}.
\] (56)

4.5 Distance to \( n \)th neighbor from the origin

Distribution of the distance to the \( n \)th neighbor from an arbitrarily chosen location \( o \) in a population of iud points distributed in a set \( W \) of arbitrary shape has been recently addressed by Srinivasa and Haenggi [16]. Particularly, they demonstrated that if \( N \) points are iud in \( W \) the pdf of the distance to the \( n \)th neighbor, \( n = 1, 2, \ldots, N \) from \( o \) follows a generalized beta distribution. We will briefly
summarize their findings below. Notice that this question in $W \subset \mathbb{R}^2$ of circular shape has also been studied by Tseng et. al [31] in 2006.

Let $R$ be the radius of a $M$-dimensional ball centered at the origin $o \in W$ and let $N$ points be iud in $W$. Assume that the center of this ball does not coincides with a point of the process. Notice that when $o$ coincides with a point of a process then the process reduces to $(N - 1)$ points iud in $W$ and the following results are still correct given that $N$ is replaced by $(N - 1)$. Let further $r_n, n = 1, 2, \ldots, N$ be Euclidian distance from $o$ to the $n$th nearest point. Let also $A_M(o, r)$ be the $M$-dimensional ball or radius $r$ centered at $o$. Fig. 3 illustrates the concept for $M = 2$.

The CDF expressing the probability that there are less than $n$ points in $A_M(o, r)$ is given by

$$1 - F(r, n) = \sum_{k=0}^{n-1} \binom{N}{k} p^k (1 - p)^{N-k}, \quad 0 \leq r \leq R. \quad (57)$$

where $F(r, n)$ is the complementary CDF and $p = |A_M(o, r) \cap W|/|W|$. The pdf of the distance function is then

$$f(r, n) = -\frac{d(1 - F(r, n))}{dr} = \frac{dp}{d} \frac{(1 - p)^{N-n} p^{n-1}}{B(N-n+1, n)}, \quad (58)$$

where $B(a, b) = \Gamma(a)\Gamma(b)/\Gamma(a + b)$ is the beta function.

The volume of a $M$-dimensional ball is $c_M R^M$, where

$$c_M = |A_M(o, 1)| = \frac{\pi^{M/2}}{\Gamma(1 + M/2)}, \quad (59)$$

is the volume of the unit ball in $\mathbb{R}^M$ [2]. Special cases are $c_1 = 2$, $c_2 = \pi$, $c_3 = 4\pi/3$. The intensity of this process is $N/c_M R^M$. 20
Recalling that \( o \) is the center of \( A_M(o, r) \) we have \( p = c_Mr^M/c_MR^M \). Using (59) we get

\[
\begin{align*}
\frac{f(r, n)}{R} &= \frac{d}{(r/R)^{M-1} (1-p)^{N-n} p^{n-1}} \frac{B(N-n+1, n)}{B(N-n+1, n)}, \\
&= \frac{d (1-p)^{N-n} p^{(n-1)/M}}{B(N-n+1, n)}, \\
&= \frac{d B(n-1/M + 1, N-n + 1N-n+1)}{B(N-n+1, n)} \times \\
&\quad \times \beta \left( \left( \frac{r}{R} \right)^M, n - \frac{1}{M} + 1, N - n + 1 \right),
\end{align*}
\]

which is generalized beta distribution defined in \( 0 \leq r \leq R \).

Authors in [16] provided a number of corollaries of this general result. For \( M = 2 \) (60) degenerates to

\[
\begin{align*}
\frac{f(r, n)}{R} &= \frac{2 \Gamma(n + 1/2) \Gamma(N + 1)}{\Gamma(n) \Gamma(N + 3/2)} \beta \left( \frac{r^2}{R^2}, n + \frac{1}{2}, N - n + 1 \right).
\end{align*}
\]

The nearest and the most distant point distances are

\[
\begin{align*}
f(r, 1) &= \frac{dN}{r} \left[ 1 - \left( \frac{r}{R} \right)^M \right]^{N-1} \left( \frac{r}{R} \right)^M, \\
f(r, N) &= \frac{dN}{r} \left( \frac{r}{R} \right)^{NM}. 
\end{align*}
\]

For one-dimensional process \( f(r, n) = f(R - r, N - n + 1) \). Thus, knowledge of the distance pdf for the nearest \( N/2 \) nodes gives complete information on the distance distributions to other points. If a point of the process coincides with arbitrarily chosen location distance distributions still follow (60) with \( N \) replaced by \( (N - 1) \). Moments can be directly obtained from (60). Finally, Haenggi compared distance distributions for \( N \) iud points in a region with that of the homogeneous Poisson process with the same intensity conditioned on the event of at least \( N \) points in a region. He showed that the latter provides rough approximation for distance distributions for the most distant neighbors. See [16] for details.

Haenggi also obtained the distance to the \( n \)th nearest neighbor for \( N \) iud points in \( l \)-sided regular polygon \( W \subset \mathbb{R}^2 \). Let \( o \) be the center of \( W \) and let \( R_i \) and \( R_c \) be inradius and circumradius of a polygon as shown in Fig. 4. They are given by

\[
\begin{align*}
R_i &= \sqrt{\frac{|W|}{l} \cot \left( \frac{\pi}{L} \right)}, \\
R_c &= \sqrt{\frac{2|W|}{l} \csc \left( \frac{2\pi}{L} \right)},
\end{align*}
\]

21
where $|W|$ is the volume of the polygon.

When $r \leq R_i$, $A_2(o, r)$ is completely within the polygon and the number of points contained in it is binomially distributed with parameters $n = N, p = \pi r^2/2$. When $R_i < r \leq R_c$, $|W \cap A_2(o, r)|$ is evaluated considering the regions of a circle lying outside the polygon. In this case the number of points follows binomial distribution with parameters $n = N$ and $q = (\pi r^2 - \pi^2 \theta + lR_i\sqrt{r^2 - R_i^2})/|W|$, where $\theta = \cos^{-1}(R_i/r)$ (see Fig. 4).

**Figure 4: Various characteristics of a polygon.**

Assuming that the polygon is centered at $o$, $|W| = A$ and no points of the process are at the origin, using (58) pdf of the distance to the $n$th neighbor is

$$f(r, n) = \begin{cases} 
\frac{2\pi r (1 - p)^{N-n}p^{n-1}}{|W| B(N - n + 1, n)}, & 0 < r \leq R_i, \\
\frac{2r(\pi - \theta) (1 - q)^{N-n}q^{n-1}}{|W| B(N - n + 1, n)}, & R_i < r \leq R_c, \\
0, & R_c < r. 
\end{cases}$$

(64)

### 4.6 Joint distance distribution to $N - 1$ neighbors

To the best of our knowledge up to date only Miller and Tseng et. al addressed joint distribution of distances from a common reference node to $N - 1$ neighbors [32, 31]. Here, we follow Miller. Given $N$ points i.i.d in a square he studied a conditional distribution of distances to other $(N - 1)$ points from a randomly chosen point located at $(x, y)$. Consider $N$ points randomly distributed in a $D \times D$ square. Assume that a square is located as shown in Fig. 5. The joint CDF of the distances between a randomly chosen reference point and the other $(N - 1)$ points is given by

$$F(\alpha_1, \ldots, \alpha_{N-1}) = \frac{8}{D^2} \int_0^{D/2} dx \int_0^x dy \prod_{n=1}^{N-1} F(\alpha_n|(x, y)),$$

(65)
where $F(\alpha_n|x, y)$ is the conditional CDF of a single point’s distance from the reference point given the position of the reference point $(x, y)$.

![Diagram](image)

Figure 5: Position of a rectangle and various angles.

Assuming that $0 \leq y \leq x \leq D$ Miller decomposed the conditional CDF as

$$F(\alpha|(x, y)) = \frac{1}{2D^2} \left( \int_{-\theta_{12}}^{\theta_{11}} d\theta \left[ \min \left( \alpha, \frac{D/2 - x}{\cos \theta} \right) \right]^2 + \right.$$  

$$\left. + \int_{-\theta_{22}}^{\theta_{21}} d\theta \left[ \min \left( \alpha, \frac{D/2 - y}{\cos \theta} \right) \right]^2 + \right.$$  

$$\left. + \int_{-\theta_{42}}^{\theta_{41}} d\theta \left[ \min \left( \alpha, \frac{D/2 + y}{\cos \theta} \right) \right]^2 + \right.$$  

$$\left. + \int_{-\theta_{12}}^{\theta_{11}} d\theta \left[ \min \left( \alpha, \frac{D/2 + x}{\cos \theta} \right) \right]^2 \right), \quad (66)$$

where the angles are given by (see Fig. 5)

$$\theta_{11} = \tan^{-1} \left( \frac{D/2 - y}{D/2 - x} \right), \quad \theta_{12} = \tan^{-1} \left( \frac{D/2 + y}{D/2 - x} \right),$$

$$\theta_{31} = \tan^{-1} \left( \frac{D/2 + y}{D/2 + x} \right), \quad \theta_{32} = \tan^{-1} \left( \frac{D/2 - y}{D/2 + x} \right),$$

$$\theta_{21} = \frac{\pi}{2} - \theta_{32}, \quad \theta_{22} = \frac{\pi}{2} - \theta_{11},$$

$$\theta_{41} = \frac{\pi}{2} - \theta_{12}, \quad \theta_{42} = \frac{\pi}{2} - \theta_{31}. \quad (67)$$

Assume that the reference point is located in the first quadrant $D/2 - x \leq D/2 - y \leq D/2 + y \leq D/2 + x$. Thus, for $\alpha \leq D/2 - x$ the integrands in (66) is
\( \alpha^2 \) and the conditional probability is \( \pi \alpha^2 / D^2 \). When \( \alpha \) increases the probability increases but is less than \( \pi \alpha^2 / D^2 \) as the integrands in (66) change from \( \alpha^2 \) to \( \beta^2 / \cos^2 \theta \) according to the following rules

\[
\min \left( \alpha, \frac{\beta}{\cos \theta} \right) = \begin{cases} 
\alpha, & \alpha < \beta, \\
\alpha, & \alpha \geq \beta, \cos \theta > \beta / \alpha, \\
\frac{\beta}{\cos \theta}, & \alpha \geq \beta, \cos \theta \leq \beta / \alpha,
\end{cases}
\]

where \( \beta \) is the shortest distance to the edge of the square in a given region.

Each of the integrals in (66) can be further decomposed as follows

\[
I = \frac{1}{2D^2} \int_{-\theta_2}^{\theta_1} d\theta \left[ \min \left( \alpha, \frac{\beta}{\cos \theta} \right) \right]^2 = \\
= \frac{\alpha^2}{2D^2} \left[ \theta_2 + \theta_1 - \min(\theta_2, \theta_1) - \min(\theta_3, \theta_2) \right] + \\
+ \frac{\beta^2}{2D^2} \left[ \tan(\min(\theta_3, \theta_1)) + \tan(\min(\theta_3, \theta_2)) \right],
\]

where \( \theta_3 = \cos^{-1}(\min(\beta, \alpha)/\alpha) \).

### 4.7 Discussion

Summarizing this section one may notice that general results for iud point patterns in a region are not available. For example, joint distribution of the distance from a randomly chosen location to the \( n \)th neighbor is not available in the closed form. At the same time one-dimensional distributions are available in rather simple form. The reason behind complex results for joint distributions is dependence in points’ occurrence in disjoint sets. It means that we no longer benefit from memoryless property deriving characterizing distance between points. This observation has two important consequences. Firstly, point processes of more complex structure would never result in manageable expressions for inter-point distance distributions. Secondly, in applied studies the Poisson process is often preferable compared to iud points in a region. Taking into account that iud points in a region naturally happen as a result of conditioning of the Poisson process one may expect that distance distributions for these two processes are close to each other.

Comparison of distance distribution for two processes has been performed in [23] and [16]. In [23] Kostin compared the distance between two neighbors in for \( N \) points distributed in a square \( A \) to the distance between nearest neighbors in homogenous Poisson process with intensity \( N|A| \), where \( |A| \) is the area of the square. Although only visual comparison has been performed it is obvious that
approximation becomes better as $N$ increases or $|A|$ decreases. Visual comparison of distances to the $n$th neighbor has been performed in [16]. Authors observed that even when $N$ is large approximation for large $n$ is unacceptable. However, it rarely happens in applied studies that the distance to the fairest neighbor is of interest. Nevertheless, these observations need to be taken into when approximation by the Poisson process is used.

5 Applications

In this section we consider few representative applications of those results summarized in the previous sections. We stress that in most applications reviewed below distance distributions between various entities are by no means a central part of studies. However, their knowledge provide the basis for analysis and without it those studies would have never been possible. At the end we will also provide few references to additional sources of information.

5.1 Coverage around a node

Consider a Poisson process in $\mathbb{R}^2$ and let us tag an arbitrary node. Assume that this node has a message to broadcast to its neighbors and there are $n$ attempts to do that. Here, we are interested in the distribution of the number of stations that receive this message after $i$th retransmission attempts. This problem is related to the optimal broadcasting and routing in ad-hoc and sensor networks as it characterize one-hop progress in the message delivery process. For $n = 1$ and Nakagami-$m$ fading this problem has been solved in [33], where the authors demonstrated that it can be reduced to probabilistic thinning of the Poisson process. The resulting process is non-homogenous as its intensity decreases with the distance. Observing that packet receptions at different nodes are independent of each other these results can be extended to $n > 1$.

5.2 Coverage problems

Another representative problem is coverage of an area when all nodes operate in broadcasting mode. Assume that communication range of each station is $r$ and all nodes in this coverage receives a message with probability 1. Nodes that are outside cannot correctly receive a message. Once again let us tag an arbitrary node. The question here is what is the area that can be covered when all stations retransmit a message. When time constraints are not introduced this problem is referred to as continuum percolation problem or, equivalently, a Boolean coverage model, see e.g. [34]. There are a number of practical applications of such model
including $k$-coverage problem is sensor networks, that is, estimating a fraction of area of the region in $\mathbb{R}^2$ or $A \in \mathbb{R}^2$ that is covered by at least $k$ sensors. Using basic principles of integral geometry coverage problems have been addressed in [35, 36] among others. To the best of our knowledge one of the most general solutions for $k$-coverage problem in sensor networks has been published in [36], where the authors assumed that sensing capabilities are not necessarily the same and used kinematic density to solve the problem.

Another practical application is finding the number of relay stations in cellular system that help a given originator to deliver a message to the base station. A special case of interest arises when the number of retransmissions is limited to a certain number. To solve this problem at each retransmission attempt coverage area of stations having the same message is to be found. Although no studies on this problem have been published yet it, this problem can be solved similarly to the coverage problem in sensor networks. Notice that this problem is also related to the bombing problem extensively studied in 50s-60s, [37] and [2], Ch.3.

Intrusion detection problems in sensor networking are also related to the coverage of the space. In order for a sensor network to detect intrusion along a certain border we should have this border fully covered by sensing ranges of sensors. Assuming a border of circular area and similar sensing ranges of sensors this problem can be reduced to the coverage of a circle by arcs. This problem has been extensively studied in literature. Among others one could refer to [13], Ch.4, [38]. For a given density of nodes one could find the probability that the circumference is fully covered by sensing ranges or probability that there will be exactly $l$ gaps, [39, 40]. Recently, in context of border intrusion detection this problem has been addressed in [41], where the authors provided extensive overview of literature. In stochastic geometry this problem is known as Boolean coverage in one dimension. In queuing theory it is related to the busy period in M/G/$\infty$ system.

A similar border coverage problem has been solved in [42], where the authors considered the probability of the coverage of a path along a straight line in a sensor networks. Using simple assumption of Poisson distribution of nodes in $\mathbb{R}^2$ this problem naturally reduces to the Boolean problem in one dimension.

### 5.3 Information dissemination

Information dissemination is another problem of interest in wireless sensor and ad-hoc networks. There are a number of particular applications of the problem. For example, one could be interested in how far a virus could propagate in a given configuration of a sensor network. Another related problem is propagation of a broadcast message generated by a certain node. For both cases, an appropriate first-order approximation can be provided using epidemic models. Once could refer to [43] for rather condensed introduction into classic epidemic models and
Application of classic epidemic models presume that the population is fully mixed, i.e. any node can reach any other node. This is very far from reality in wireless multi-hop networks. When message broadcasting is of interest, those nodes that received a message and have already successfully sent it to all the neighbors no longer contribute in the message propagation process. Thus, only rough approximation is provided using classic epidemic models. Recently, epidemic models on structured population have been introduced. According to such models a graph of a network is constructed first and then percolation theory is used to estimate performance metrics of interest, e.g. number of stations having a message after some time, coverage of the whole network, etc. In order to construct a graph, distance distributions between nodes in a network are needed. See [45] for concise introduction to epidemics on structured populations and [46] for applications to the message propagation in sensor networks.

5.4 Interference

Description of the amount of interference at a certain (possibly random) location in a network created by a number of transmitting stations is another application of distance distributions in wireless networks. This question is often of high importance in context of planning of cellular systems when one is choosing an appropriate frequency reuse plan. It is also a question of interest in wireless ad-hoc and sensor networks where a number of stations may operate using the same channel in different spatial regions of a network.

Consider an arbitrary point $x$ and a number of stations iid in $W$ of circular area (cellular systems) or distributed according to homogenous Poisson process in $\mathbb{R}^2$ (ad-hoc or sensor networks). This problem of interference at $x$ can be solved by summing up contributions of all nodes relative to their distances from $x$. For detailed analysis one could refer to [47] or [48] Section III for short overview.

5.5 Routing in ad-hoc networks

Yet another issue is routing in wireless ad-hoc and sensor systems. In order to ensure efficient routing progress needs to be made at each hop. To characterize it we need to estimate the distance to the nearest neighbor within an angle $0 < \phi \leq \pi/2$ at each relaying attempt. According to [10], in terms of distribution this correspond to the change of a volume from $M$-dimensional ball to $M$-dimensional sector with open angle. Its volume is $c_{\phi,M} r^M$, i.e., $c_{\phi,1} = 1$, $c_{\phi,2} = \phi$, $c_{\phi,3} = 2\pi(1 - \cos \phi)]/3$. Then, using (58) the distribution of Euclidian distance to the
nth neighbor in a sector $\phi$ is

$$f(r, n) = \frac{M(\lambda c, M r^M)^n}{r \Gamma(n)} e^{-\lambda c, M r^M}, \quad r \geq 0, n = 1, 2, \ldots$$  \hspace{1cm} (70)$$

Using (70) one could estimate the progress to the next hop at each relaying point. Notice that for $M = 2$, $\phi = \pi/4$, $r_1$ has Rayleigh distribution. For more details, e.g. moments, see [10]. For treatment of the same problem for $N$ iud nodes in a region see [16].

5.6 Further references

The abovementioned review of applications of distance distributions in random networks is by no means exhaustive. A number of additional use cases are highlighted in excellent review of stochastic geometry’s application to wireless networks [48]. A sequence of studies by Haenggi illustrates the usage of distance distributions between points of a Poisson process in $\mathbb{R}^2$ and more recently iud points in a region in various problems of communications in multi-hop relay networks, see e.g. [10, 16] for details.

6 Conclusions

In this paper we surveyed results for various distance distributions in two mostly used spatial models - spatial Poisson process and fixed number of iud points in a spatial region. One of the most interesting observations is that even for such simple spatial models results for distance distributions can be very complicated to derive. As one expects results for fixed number of iud points in a regions are also more complex compared to those of homogenous Poisson process in $\mathbb{R}^2$. An important consequence is that for more complicated and possibly more realistic models distance distributions between points would be impossible to obtain. Still simple models considered in this paper do allow for closed-form expressions of important distance distributions and may serve as first-order approximation for real nodes’ distribution. We also would like to note that although most presented results concern $\mathbb{R}^2$ as a natural space for networking applications some of them can be extended to $\mathbb{R}^M$.

We also briefly reviewed application areas of distance distributions in cellular, ad-hoc and sensor networks including connectivity, coverage of the space, information dissemination, interference and routing problems. We saw that very often spatial point processes serve as an underlying structure for more complex models, e.g. Boolean coverage models. In this case availability of various distance distributions in a closed form is a must for proper assessment of the child models.
References


