The Distribution of the Size of the Intersection of a $k$-Tuple of Intervals

Vladimir Bozovic, Shanzhen Gao, and Heinrich Niederhausen
Florida Atlantic University, Boca Raton

Abstract

Let $(I_1, I_2, ..., I_k)$ be a random $k$-tuple of subintervals of the discrete interval $[1, n]$, and $L_n$ the random variable that measures the size of their intersection. We derive the exact and asymptotic distribution of $L_n$ under the assumption of equally likely drawn $k$-tuples. The enumeration of such $k$-tuples and refinements of the given statistic lead to interesting relations to other topics, like octahedral numbers and bipartite graphs.

1 Introduction

How do we decide whether a sequence of subintervals from $[1, n] = \{1, \ldots, n\}$ is “random”, i.e., independently and equally likely drawn from all $\left(\frac{n+1}{2}\right)^k$ subintervals, if all we get to see is the size of their intersection? This paper was motivated by investigating the distribution of the intersection size of $k$ subintervals of $[1, n]$. For example, if 5 subintervals from $[1, 10]$ intersect in 4 or more points, we can be 99.5% certain that they are not randomly generated. Alternatively, suppose we can draw the subintervals one after the other. If $n$ is larger than 4, and the first 10 intervals still have a nonempty intersection, we can be more than 99% sure that the intervals are not random. We will see in Section 6 that the probability of drawing $k$ intersecting intervals approaches $2^k / \binom{2k}{k} \approx 2^{-k} \sqrt{k} \pi$ for $n \to \infty$.

Let $\Lambda_{n,l} := \{(I_1, I_2, ..., I_k) \mid I_j$ is a subinterval of $[1, n]$ for all $j \in [1, k]$, and $\bigcap_{j=1}^{k} I_j = l\}$. The enumeration begins with the basic observation in Section 2 that $\Lambda_{n,l} = \Lambda_{n+1-l,1}$ for $l > 0$. Whenever feasible, we will therefore write $\Lambda_j^k$ instead of $\Lambda_{j;1}^k$. In this notation

$$
\Pr(\text{The size of the intersection of 5 subintervals of } [1, 10] \text{ is } 4 \text{ or larger})
= \left(\frac{11}{2}\right)^{-5} \sum_{l=4}^{10} |\Lambda_{11-l,1}^5| = 0.00474
$$

using one of the many formulas we will derive for $|\Lambda_{n,1}^k|$. 
A combinatorial interpretation of the coefficients (called $c(p,k)$) of $\binom{n}{k}$ in the expansion of $\Lambda_k^n$ will be given in Section 4.1. We found the (octahedral) numbers $|\Lambda^2_n|$ especially noteworthy; they are discussed in Section 3.

Expanding $|\Lambda^k_n|$ in powers of $n$ shows another interesting feature:

$$|\Lambda^2_n| = \frac{1}{3} n + \frac{2}{3} n^3, \quad |\Lambda^3_n| = \frac{1}{2} n + \frac{1}{2} n^3 + \frac{3}{5} n^5,$$

$$|\Lambda^4_n| = \frac{4}{35} n^7 + \frac{4}{35} n^7 + \frac{2}{5} n^5 + \frac{23}{105} n, \quad |\Lambda^5_n| = \frac{2}{7} n - \frac{5}{126} n^3 + \frac{1}{3} n^5 + \frac{5}{21} n^7 + \frac{5}{126} n^9.$$

Only odd powers of $n$ occur in those expansion; however, the negative coefficient in $|\Lambda^5_n|$ discourages a search for combinatorial significance. It turns out that the Bernoulli numbers $B_k$ are to blame (Section 5); on the other hand, they give us a rather precise approximation of our numbers, $|\Lambda^k_n| \approx \frac{\left(\binom{n+1}{2k+1}-2n^{2k+1}+(n-1)^{2k+1}\right) k!}{(2k+1)!} - \frac{2^{(n+1)^k-2(n^k+(n-1)^k)}B_{k+1}}{k+1}$ for odd $k$, and a similar approximation for even $k$ (Corollary 9). The numerical experiments hint at another expansion again in odd degrees but of a different basis,

$$|\Lambda^2_n| = \binom{n}{1} + 4\binom{n+1}{3}, \quad |\Lambda^3_n| = \binom{n}{1} + 12\binom{n+1}{5} + 36\binom{n+2}{5},$$

$$|\Lambda^4_n| = \binom{n}{1} + 28\binom{n+1}{7} + 240\binom{n+2}{7} + 576\binom{n+3}{7}.$$

Those coefficients are indeed positive integers, and they are derived in Section 5.1. The most detailed refinement of $\Lambda^k_n$ that we consider is the number of all $k$-tuples that intersect in the single number $p$, consist of $h$ different subintervals, and have $u$ left endpoints and $v$ right endpoints. The number of such $k$-tuple of subintervals equals \(\binom{p-1}{w-1}\binom{n-p}{v-1}\) $B(u, v, h)S(k, h)!$, (Section 4.3), where $S(k, h)$ stands for the Stirling numbers of the second kind, and $B(u, v, h)$ is the number of ways to select $h$ elements from an $u \times v$ matrix such that at least one element
is chosen from each row and each column. At the same time, the numbers $B(u, v, h)$ are the connection coefficients in the product formula

$$
\binom{mn}{h} = \sum_{i=0}^{m} \sum_{j=0}^{n} \binom{m}{i} \binom{n}{j} B(i, j, h)
$$

bringing a combinatorial interpretation to a rather bland application of linear algebra (Remark 6). There is a closely related second interpretation of this formula in terms of bipartite graphs (Remark 5).

Deriving the size of the intersection is a classical topic in computer science; an algorithm was presented by McCreight [3]. For most of our expansions we need properties of Stirling numbers. Jacobo Stirling published the first table of “his” numbers in 1730, and a systematic treatment of Stirling numbers appeared in Jordan’s Calculus of Differences [2] 200 years later. However, for convenience and accessibility we will refer to R. Stanley’s text book, Enumerative Combinatorics I, whenever possible.

2 Two Basic Observations

The following two simple lemmas are the basis for our combinatorial approaches.

**Lemma 1** $\Lambda^k_{p,l} = \Lambda^k_{n+1-l;1}$ for $l > 0$

**Proof.** It suffices to show that $\Lambda^k_{n,l} = \Lambda^k_{n-1,l-1}$ for $l > 1$. Suppose $(I_1, I_2, ..., I_k)$ is a $k$-tupel of subintervals of $[1, n]$, and $\bigcap_{j=1}^{k} I_j = l$. The intersection is again an interval, and let $r$ be the right endpoint of this interval. Now remove $r$ from each of the $k$ intervals, and decrease all elements larger than $r$ by 1. This way all $k$ intervals are mapped into subintervals of $[1, n-1]$, and their intersection has size $l - 1$. Vice versa, we can any $k$-tupel from $\Lambda^k_{n-1,l-1}$ to a $k$-tupel by increasing all numbers larger than the right endpoint of the $l - 1$-element intersection interval by 1, and enlarging the intersection interval by 1 at the right end.

Sequences of subintervals from $[1, n]$ are nothing but sequences of ordered pairs of endpoints. The following lemma translates the intersection cardinality into a condition on the endpoints.

**Lemma 2** Let $I_j = [l_j, r_j], j = 1, 2, ..., k$ be intervals such that $I_j \subseteq [1, n]$. Then

$$
x \in \bigcap_{j=1}^{k} I_j \Leftrightarrow \max_{1 \leq i \leq k} \{l_j\} \leq x \leq \min_{1 \leq j \leq k} \{r_j\}
$$

Furthermore,

$$
\big| \bigcap_{j=1}^{k} I_j \big| = l \Leftrightarrow \min_{1 \leq j \leq k} \{r_j\} - \max_{1 \leq i \leq k} \{l_j\} = l - 1
$$
**Proof.** The lemma follows because

\[ x \in \bigcap_{j=1}^{k} I_j \iff l_j \leq x \leq r_j, \quad j = 1, \ldots, k \iff \max_{1 \leq i \leq k} \{l_i\} \leq x \leq \min_{1 \leq j \leq k} \{r_j\} \]

\[\blacksquare\]

### 3 Octahedral Numbers

Let \( \Lambda_n^2 \) be the set of pairs of discrete subintervals of \([1, n]\) that intersect in one point. If \( I = [a, b] \) is the first and \( J = [c, d] \) the second interval, then we have following situations for which \( |I \cap J| = 1 \),

\[
\begin{align*}
    a &= b = c = d \\
    a &= b = c < d, \quad c = d = a < b, a < b = c = d, c < d = a = b \\
    a < b = c < d, c < d = a < b, a < c = d < b, c < a = b < d
\end{align*}
\]

Considering the number of different points (1, 2 or 3) in the above cases it is easy to see that

\[ |\Lambda_n^2| = \binom{n}{1} + 4\binom{n}{2} + 4\binom{n}{3} \]

If we rearrange the expression on the right hand side, we get the well known octahedral numbers

\[ 1 + 2^2 + \ldots + (n - 1)^2 + n^2 + (n - 1)^2 + \ldots + 2^2 + 1 \]

A bijection between \( \Lambda_n^2 \) and the discrete octahedron with an \( n \times n \) center square can be obtained as follows. Let \( x \wedge y := \min \{x, y\} \). In the Euclidean 3-space \( \mathbb{Z}^3 \) we represent the octahedron as a union of squares (layers)

\[ A_l = \{(x, y, l) \mid ||l|| < (x \wedge y) \leq n, \quad x, y \in \mathbb{N}\} \]

for \( l \in [-n + 1, n - 1] \).
Each layer $A_l$ contains $(n - |l|)^2$ points with integral coordinates. Therefore,
\[
\sum_{l=-n+1}^{n+1} |A_l| = 1 + 2^2 + \ldots + (n-1)^2 + n^2 + (n-1)^2 + \ldots + 2^2 + 1
\]

Consider the mapping $\Theta : \bigcup_{l=-n+1}^{n-1} A_l \rightarrow \Lambda_n^2$ defined as
\[
\Theta(x, y, l) = \begin{cases} 
([x \land y, x], [l, y]) & \text{if } l < 0 \\
([x \land y, y], [x \land y, x]) & \text{if } l = 0 \\
([l, y], [x \land y, x]) & \text{if } l > 0 
\end{cases}
\]

Obviously
\[
\Theta(x, y, l) = \Theta(x_1, y_1, l) \Rightarrow (x, y) = (x_1, y_1)
\]
for all $l \in [-n + 1, n - 1]$. Hence $\Theta$ is injective. Consequently, define $\Theta^{-1} : \Lambda_n^2 \rightarrow \bigcup_{l=-n+1}^{n-1} A_l$ by
\[
\Theta^{-1}([l_1, r_1], [l_2, r_2]) = \begin{cases} 
(r_1, r_2, -l_2) & \text{if } l_2 < l_1 \\
(r_1, r_2, 0) & \text{if } l_2 = l_1 \\
(r_2, r_1, l_1) & \text{if } l_1 < l_2 
\end{cases}
\]

Again, it is easy to check that $\Theta^{-1}$ is injective. We show that $\Theta$ is a bijection by proving that $\Theta^{-1}$ is indeed its inverse. Let $(x, y, l) \in A_l$, thus $|l| < (x \land y) \leq n$.
If $l > 0$, then $\Theta^{-1}(\Theta((x, y, l))) = \Theta^{-1}([l, y], [x \land y, x]) = (x, y, l)$.
If $l < 0$ then $\Theta^{-1}(\Theta(x, y, l)) = \Theta^{-1}([x \land y, x], [l, y]) = (x, y, -|l|)$.
If $l = 0$ then $\Theta^{-1}(\Theta(x, y, 0)) = \Theta^{-1}([x \land y, x], [x \land y, y]) = (x, y, 0)$.

4 The General Case

From Lemma 2 follows that $(I_1, I_2, \ldots, I_k) \in \Lambda_n^k$ iff
\[
\max_{1 \leq j \leq k} \{l_j\} = \min_{1 \leq j \leq k} \{r_j\}
\]
where $l_j$ are the left endpoints and $r_j$ are the right endpoints of the intervals $I_j = [l_j, r_j]$ in this $k$-tuple. Suppose the intervals intersect in $p \in [1, n]$. The number of ways the left endpoint can be chosen equals the number $p^k - p^{k-1}$ of mappings from $[1, k]$ to $[1, p]$ that contain $p$ as an image. Interpreting the right endpoints in the same way shows that
\[
|\Lambda_n^k| = \sum_{p=1}^{n} \left( p^k - (p-1)^k \right) \left( (n+1-p)^k - (n-p)^k \right)
\]

This is probably the most “basic” answer to our problem. It hides the polynomial character of the numbers $|\Lambda_n^k|$, which will become more apparent in the following refinements of the problem.
Note that
\[
\sum_{l=1}^{n} |\Lambda^k_l| = \sum_{p=1}^{n} \left( p^k - (p - 1)^k \right) \sum_{l=p}^{n} \left( (l + 1 - p)^k - (l - p)^k \right) = \sum_{p=1}^{n} \left( p^k - (p - 1)^k \right) (n + 1 - p)^k
\] (2)

This sum plays a role in determining the probability of selecting \( k \) nonintersecting subintervals (Section 6).

4.1 Endpoint Sets

Consider the mapping \( \Phi : \Lambda^k_n \to [1, n] \) defined as
\[
\Phi([l_1, r_1], [l_2, r_2], ..., [l_k, r_k]) = \bigcup_{i=1}^{k} \{l_i, r_i\}.
\]
The set on the right-hand side in the argument of the mapping \( \Phi \) we will call the endpoint set of the \( k \)-tuple. We know from Lemma 2 that
\[
\Phi([l_1, r_1], [l_2, r_2], ..., [l_k, r_k]) = [1, m].
\]
and we can divide set \( \Lambda^k_n \) into equivalence classes regarding different endpoint sets. We are interested in the sizes of those equivalence classes, i.e., we want to find \( |\Phi^{-1}\{i_1, i_2, ..., i_m\}| \) where \( \{i_1, i_2, ..., i_m\} \subseteq [1, n] \) for some fixed \( m \), such that \( 1 \leq m \leq 2k - 1 \). Without loss of generality we can assume that \( \{i_1, i_2, ..., i_m\} = [1, m] \). Therefore, we want the number of all \( k \)-tuples from \( \Lambda^k_n \) such that
\[
\Phi([l_1, r_1], [l_2, r_2], ..., [l_k, r_k]) = [1, m].
\]
Denote this number as \( c(m, k) = |\Phi^{-1}[1, m]| \). For finding this number it is helpful to notice that
- every number from \([1, m]\) has to occur in the corresponding \( k \)-tuple \( ([l_1, r_1], [l_2, r_2], ..., [l_k, r_k]) \in [1, m] \) at least once.
- since \( k \)-tuples from \( \Phi^{-1}[1, m] \) intersect in one point, that intersection point must be in the set \([1, m]\).

Suppose that \( \bigcap_{j=1}^{k} [l_j, r_j] = \{p\} \), where \( p \in [1, m] \). From \( \max_{1 \leq j \leq k} \{l_j\} = p = \min_{1 \leq j \leq k} \{r_j\} \) (Lemma 2) follows \( \{l_1, l_2, ..., l_k\} = [1, p] \) and \( \{r_1, r_2, ..., r_k\} = [p, m] \). Denote the number of occurrences of \( i \in [1, p] \) among the left end \( k \)-tuple \( (l_1, l_2, ..., l_k) \) as \( t_i \). Therefore, \( t_1 + t_2 + ... + t_p = k \) where \( t_i \geq 1 \) for all \( i \in [1, p] \). For every such composition \( t_1 + t_2 + ... + t_p = k \) we have \( \binom{k}{t_1, t_2, ..., t_p} \) different orders of left endpoints. Similarly, we have \( \binom{k}{w_p, w_p+1, ..., w_m} \) different
orders of right endpoints for the composition \( w_p + w_{p+1} + \ldots + w_m = k \), where \( w_i \) represents the number of occurrences of \( i \in [p, m] \) among the right endpoint \( k \)-tuple \( (r_1, r_2, \ldots, r_k) \). Thus

\[
c(m, k) = \sum_{p=1}^{m} \sum_{t_u \geq 1, t_1 + t_2 + \ldots + t_p = k} \binom{k}{t_1, t_2, \ldots, t_p} \times \sum_{w_v \geq 1, w_p + w_{p+1} + \ldots + w_m = k} \binom{k}{w_p, w_{p+1}, \ldots, w_m}
\]

It is well known [2 § 60] that

\[
\sum_{t_u \geq 1, t_1 + t_2 + \ldots + t_p = k} \binom{k}{t_1, t_2, \ldots, t_p} = p! S(k, p) = \sum_{j=0}^{p} \binom{p}{j} (-1)^{p-j} j^k,
\]

the Stirling number of the second kind. Indeed, this may be seen as a direct consequence of the exponential generating function of the Stirling numbers of the second kind (as in Jordan’s book), or as an application of the multinomial formula and the exclusion-inclusion principle. Hence

\[
c(m, k) = \sum_{p=1}^{m} p! S(k, p)(m + 1 - p)! S(k, m + 1 - p) \quad (3)
\]

Now we can easily evaluate the number of all \( k \)-tuples from the set \( \Lambda_n^k \) that intersect in \( m \) as \( \binom{n}{m} c(m, k) \). As we noticed earlier, \( 1 \leq m \leq 2k - 1 \), hence

\[
|\Lambda_n^k| = \sum_{m=1}^{2k-1} \binom{n}{m} c(m, k) = \sum_{m=1}^{2k-1} \binom{n}{m} \sum_{p=1}^{m} p! S(k, p)(m + 1 - p)! S(k, m + 1 - p) \quad (4)
\]

This formula for \( |\Lambda_n^k| \) makes it obvious that \( |\Lambda_n^k| \) can be extended to a polynomial \( \lambda_{2k-1}(n) \) of degree \( 2k - 1 \) in \( n \).

**Remark 3** Formula (3) motivates us to interpret the total number of different orders of left endpoints of \( k \)-tuples from \( \Phi^{-1}[1, m] \) intersecting in \( \{p\} \), as the number of surjective mappings from the set \( \{l_1, l_2, \ldots, l_k\} \) to the set \( [1, p] \). That number is equal to \( p! S(k, p) \) (see [2 1.4]), and similarly for right endpoints the number of surjections from \( \{r_1, r_2, \ldots, r_k\} \) to \( [p, m] \) equals \( (m + 1 - p)! S(k, m + 1 - p) \).

### 4.2 Geometric Interpretation

We will map intervals to points via the mapping \( F : \{(i, j) \mid [i, j] \subseteq [1, n] \} \rightarrow \{(i, j) \mid 1 \leq i \leq j \leq n \} \) defined as \( F([i, j]) = (i, j) \). Clearly, \( F \) is a bijection between the set of all subintervals of \([1, n]\) and a discrete right triangle in the coordinate plane. This bijection provides a geometric interpretation of our problem. As we pointed out earlier, every \( k \)-tuple from the set \( \Lambda_n^k \) must have an intersection in \( 1, 2, \ldots \) or \( n \). Suppose the \( k \)-tuple \( ([l_1, r_1], [l_2, r_2], \ldots, [l_k, r_k]) \) intersects in \( \{p\} \). Then we know that \( \max_{1 \leq j \leq k} l_j = p = \min_{1 \leq j \leq k} r_j \). This means that every interval \([l_i, r_i]\) from the observed \( k \)-tuple is mapped by \( F \) to the point \((l_i, r_i)\) in the rectangle \( R_p = \{(i, j) \mid 1 \leq i \leq p \leq j \leq n \} \).
Conversely, every choice of \( k \) points (with multiplicities) from the rectangle \( R_p \) containing at least one point with first coordinate equal to \( p \), and one point with second coordinate equal to \( p \), will give us a \( k \)-tuple that intersects in \( \{p\} \). Notice that the number of different points which we take from \( R_p \) in this way equals the number of different intervals in the \( k \)-tuple that intersect in \( p \). This leads us to a new way of expressing the cardinality of the set \( \Lambda^k_n \). In the following subsection we will obtain the cardinality of \( \Lambda^k_n \) by summing up the number of all \( k \)-tuples which consist of \( h \) different intervals, \( h = 1, 2, \ldots, k \).

### 4.2.1 Number of all \( k \)-tuples consisting of \( h \) different intervals

Consider the mapping \( \Gamma : \Lambda^k_n \to \mathbb{N} \) defined as

\[
\Gamma(I_1, I_2, \ldots, I_k) = |\{I_1, I_2, \ldots, I_k\}|
\]

This mapping brakes \( \Lambda^k_n \) into equivalence classes such that \( \Lambda^k_n = \bigcup_{h=1}^{k} \Gamma^{-1}(h) \), which implies \( |\Lambda^k_n| = \sum_{h=1}^{k} |\Gamma^{-1}(h)| \). Let us investigate the size of \( |\Gamma^{-1}(h)| \). As we already know, the intersection of every tuple from \( \Gamma^{-1}(h) \) must be an element of \([1, n] \). Suppose the \( k \)-tuple intersects in some number \( p \). We count the number of all \( k \)-tuples which intersect in \( p \) and consists of exactly \( h \) different intervals. Going back to the geometric interpretation, denote by \( T(p, h) \) the number of ways of choosing \( h \) different points from the rectangle \( R_p \) such that at least one is chosen with \( x \)-coordinate \( p \), and at least one with \( y \)-coordinate \( p \). It is not difficult to see, by the inclusion-exclusion principle, that:

\[
T(p, h) = \binom{p(n-p+1)}{h} - \binom{(p-1)(n-p+1)}{h} - \binom{p(n-p)}{h} + \binom{(p-1)(n-p)}{h}
\]

From \( h \) different intervals we can make \( h!S(k, h) \) different \( k \)-tuples. Therefore, the number of \( k \)-tuples from the set \( \Gamma^{-1}(h) \) which intersect in \( p \) is \( T(p, h)h!S(k, h) \). It follows that

\[
|\Gamma^{-1}(h)| = \sum_{p=1}^{n} T(p, h)h!S(k, h)
\]
and

$$|\Lambda_n^k| = \sum_{h=1}^{k} |\Gamma^{-1}(h)| = \sum_{h=1}^{k} \sum_{p=1}^{n} T(p, h) h! S(k, h) =$$

$$= \sum_{h=1}^{k} \sum_{p=1}^{n} \left( \binom{p(n-p+1)}{h} - \binom{(p-1)(n-p+1)}{h} \right) - \binom{p(n-p)}{h} + \binom{(p-1)(n-p)}{h} \right) h! S(k, h)$$

4.3 The Number of $k$-Tuples from $\Gamma^{-1}(h)$ Determined by the Number of Left and Right Endpoints

All refinements of $|\Lambda_n^k|$ so far seem to indicate that Stirling numbers of the second kind are an unavoidable component. However, if we want to calculate the number of all $k$-tuples from $\Gamma^{-1}(h)$ such that $|\{l_1, l_2, ..., l_k\}| = u$ and $|\{r_1, r_2, ..., r_k\}| = v$ then the Stirling numbers of the first kind will appear in this refinement of the formula for $|\Lambda_n^k|$. This leads us to a new approach in calculating $|\Lambda_n^k|$. We use the standard method for determining the number of $k$-tuples which satisfy some conditions, by dividing them into disjoint subsets regarding the intersection point. Therefore, for the purpose of finding all $k$-tuples from $\Gamma^{-1}(h)$ with a fixed number of left endpoints $u$ and a fixed number of right endpoints $v$, we first observe just those that intersect in a given number $p$. In our geometric interpretation this means that we are choosing $h$ different points from $R_p$ such that at least one is chosen from column $p$ (i.e., with $x$-coordinate $p$), and at least one from row $p$ (with $y$-coordinate $p$). In addition, the number of selected columns (number of left endpoints) from $R_p$ has to be exactly $u$, and the number of rows (the number of right endpoints) from $R_p$ has to be exactly $v$.

Selecting $h$ Points from $R_p$

Since we must always choose column $p$ and row $p$, we conclude that we have to choose another $u-1$ columns and $v-1$ rows. In how many different ways can this be done? The answer to this question is related to the answer of the next problem.
The disappearing matrix  Consider a matrix of dimension $i \times j$. We select $h$ elements from the matrix and erase every row and column that contain any of the $h$ elements. In how many different ways can we select $h$ elements such that the whole matrix must be erased? Denote this number of ways by $B(i, j, h)$. Notice that $B(i, j, h) = B(j, i, h)$.

We call any two sets of $h$ matrix elements from an $m \times n$ matrix equivalent, if they erase the same submatrix. If the erased submatrix has $i$ rows and $j$ columns, then the size of the corresponding equivalence class is $B(i, j, h)$. An $m \times n$-matrix has $\binom{m}{i} \binom{n}{j}$ submatrices of size $i \times j$. Hence our partitioning of the $\binom{mn}{h}$ subsets of $h$ matrix elements into equivalence classes shows that

$$\binom{mn}{h} = \sum_{i=0}^{m} \sum_{j=0}^{n} \binom{m}{i} \binom{n}{j} B(i, j, h)$$

(5)

Remark 4 Let $p_h(x) := \binom{x}{h}$, a polynomial of degree $h$ in $x$. The sequence $(p_h)_{h \geq 0}$ is an example of an important class, the polynomials of binomial type, for which the “binomial theorem” holds,

$$p_h(m + n) = \sum_{i=0}^{h} p_i(m) p_{h-i}(n)$$

Of course, $x^h / h!$ is the most prominent member of this family. However, $\binom{x}{h}$ is a “product formula”, and such formulas are usually more difficult to obtain, with one notable exception, $x^h$. We will give an algebraic proof of (5) in Remark 6 below that will uncover the actual multiplication step $(mn)^h = m^n h^h$ in this formula.

Remark 5 From the definition of $B(i, j, h)$ it is easy to see that $B(i, j, h)$ is also the number of bipartite labelled simple graphs $G([i], [j]; [h])$ without isolated vertices, vertex sets of cardinality $i$ and $j$, respectively, and edge set of size $h$. The product formula (5) gives the number of graphs $G([n], [m]; [h])$ (isolated vertices allowed). Summing over the number of edges results in the well known formula $2^{nm}$ for the number of bipartite labelled simple graphs with vertex sets of cardinality $m$ and $n$, respectively. See [4] for more general results on the enumeration of $k$-colored labelled graphs.

Let us suppose for a moment that we have an explicit form of the numbers $B(u, v, h)$. Now the problem of finding the number of all $k$-tuples from $\Gamma^{-1}(h)$ with $u$ left endpoints and $v$ right endpoints becomes much easier. Let us consider those $k$-tuples that intersect in the number $p$. As we explained before, it is enough to take $u - 1$ columns and $v - 1$ rows from $R_p$. Therefore, we select some matrix of dimension $u \times v$, and $h$ points which we take in order to “erase” the whole matrix. It follows that

$$\binom{p - 1}{u - 1} \binom{n - p}{v - 1} B(u, v, h)$$

is the number of ways of choosing $h$ points from the rectangle $R_p$ such that they take exactly $u$ columns and $v$ rows including row $p$ and column $p$. As we saw before, $h$ different intervals can be copied to $k$ places in $S(k, h) h!$ different ways, which is simply the number of surjective mappings from a $k$ element set to an $h$ element set. Finally,

$$\binom{p - 1}{u - 1} \binom{n - p}{v - 1} B(u, v, h) S(k, h) h!$$
is the number of all \( k \)-tuples that intersect in \( p \), consist of \( h \) different subintervals, have \( u \) left endpoints and \( v \) right endpoints. From this very detailed statistic on the \( k \)-tuples we can obtain many results. For example, if we sum over the intersection point \( p \), we get the number

\[
\sum_{p=1}^{n} \binom{p-1}{u-1} \frac{1}{n-p} \binom{n-p}{v-1} B(u,v,h) S(k,h) h! = \binom{n}{v+u-1} B(u,v,h) S(k,h) h! \tag{6}
\]

of all \( k \)-tuples from \( \Lambda^k_n \) that belong to \( \Gamma^{-1}(h) \) (consist of \( h \) different intervals) having \( u \) left endpoints and \( v \) right endpoints.

If we sum \( (6) \) over the number \( u \) of left endpoints, we get the number

\[
\sum_{u=1}^{k} \binom{n}{v+u-1} B(u,v,h) S(k,h) h!,
\]

of all \( k \)-tuples which belong to \( \Gamma^{-1}(h) \), and have exactly \( v \) right endpoints.

If we sum \( (6) \) over the number of left and right endpoints, we obtain the number of all \( k \)-tuples from \( \Lambda^k_n \) which belong to \( \Gamma^{-1}(h) \),

\[
|\Gamma^{-1}(h)| = \sum_{u=1}^{k} \sum_{v=1}^{k} \binom{n}{v+u-1} B(u,v,h) S(k,h) h!.
\]

If we keep the number of endpoints fixed, and sum \( (6) \) over the number \( h \) of different intervals, then we obtain the number

\[
\sum_{h=1}^{k} \binom{n}{v+u-1} B(u,v,h) S(k,h) h! \tag{7}
\]

of all \( k \)-tuples from \( \Lambda^k_n \) that have exactly \( u \) left endpoints and \( v \) right endpoints.

Finally, we find \( |\Lambda^k_n| \),

\[
|\Lambda^k_n| = \sum_{h=1}^{k} \sum_{v=1}^{k} \sum_{u=1}^{k} \binom{n}{v+u-1} B(u,v,h) S(k,h) h! \tag{8}
\]

**Evaluating** \( B(u,v,h) \)  
As we already noticed in \( (6) \), the number of all \( k \)-tuples from \( \Lambda^k_n \) that have exactly \( u \) left endpoints and \( v \) right endpoints equals

\[
\left( \binom{n}{u+v-1} \right) \sum_{h=1}^{k} B(u,v,h) S(k,h) h! \]

On the other hand, the number of all \( k \)-tuples from \( \Lambda^k_n \) that have exactly \( u \) different left endpoints and \( v \) different right endpoints can be evaluated as follows. Every \( k \)-tuple with \( u \)
left endpoints and \( v \) right endpoints takes exactly \( u + v - 1 \) numbers from \([1, n]\) (remember that left and right endpoints must overlap in one point). Therefore, we can choose \( \binom{n}{u+v-1} \) different subsets of \([1, n]\) to construct \( k \)-tuples of this type. The left endpoints take the first \( u \) places; this can be done in \( u!S(k, u) \) different ways. Similarly, there are \( v!S(k, v) \) ways to arrange the right endpoints. Therefore, the number of all \( k \)-tuples from \( \Lambda^k_n \) that have exactly \( u \) left endpoints and \( v \) right endpoints is

\[
\binom{n}{u+v-1} u!S(k, u)v!S(k, v).
\]

Comparing the last two formulas, we see that

\[
\sum_{h=1}^{k} B(u, v, h)S(k, h)h! = u!S(k, u)v!S(k, v)
\]

The well known inversion formula for Stirling numbers \([5, \text{Prop. 1.4.1}]\) shows that

\[
B(u, v, h) = \frac{u!v!}{h!} \sum_{r=0}^{h} s(h, r)S(r, u)S(r, v)
\]

where \( s(h, i) \) stands for the (alternating) Stirling numbers of the first kind.

**Remark 6** We can now verify the product formula \([5]\).

\[
\sum_{i=0}^{m} \sum_{j=0}^{n} \binom{m}{i} \binom{n}{j} B(i, j, h) = \sum_{i=0}^{m} \sum_{j=0}^{n} \binom{m}{i} \binom{n}{j} \frac{i!j!}{h!} \sum_{r=0}^{h} s(h, r)S(r, i)S(r, j)
\]

\[
= m!n! \sum_{r=0}^{h} \frac{s(h, r)}{h!} \sum_{i=0}^{m} \frac{S(r, i)}{(m-i)!} \sum_{j=0}^{n} \frac{S(r, j)}{(n-j)!}
\]

From \( \sum_{i=0}^{m} \frac{S(r, i)}{(m-i)!} = \frac{e^x - 1}{m!} \) and \( \sum_{r=0}^{h} \frac{s(h, r)}{h!} x^r = \binom{x}{h} \) (see \([5, \text{1.4}]\)) follows

\[
\sum_{i=0}^{m} \sum_{j=0}^{n} \binom{m}{i} \binom{n}{j} B(i, j, h) = \sum_{r=0}^{h} \frac{s(h, r)}{h!} (mn)^r = \binom{mn}{h}
\]

Note that this type of product formula for a polynomial sequence \((p_n(x))_{n \geq 0}\), say, is achieved by straightforward basis transformation \( p_n(x) \mapsto x^n \) and its inverse, \( x^m \mapsto p_n(x) \), in the vector space of polynomials. Such a product formula will surface whenever an inverse pair of transformations is explicitly known.
We also check our formula for $B(u, v, h)$ by calculating the cardinality of $\Lambda_n^k$ according to (8), using the identity $\delta_{i, k} = \sum_{h=1}^{k} S(k, h) s(h, i)$ ([15, 1.4.1])

$$|\Lambda_n^k| = \sum_{h=1}^{k} S(k, h) \sum_{v=1}^{k} \sum_{u=1}^{k} \left(\frac{n}{v+u-1}\right) u!v! \sum_{i=0}^{h} s(h, i) S(i, u) S(i, v)$$

$$= \sum_{v=1}^{k} \sum_{u=1}^{k} \left(\frac{n}{v+u-1}\right) u!v! S(k, u) S(k, v)$$

$$= \sum_{p=1}^{2k-1} \sum_{s=1}^{k} \left(\frac{n}{p}\right) s! (p+1-s)! S(k, s) S(k, p+1-s)$$

in agreement with (4).

5 The Polynomials

We saw in (4) that the numbers $|\Lambda_n^k|$ can be extended from their support to polynomials $\lambda_{2k-1}(x)$, say, in $x \in \mathbb{R}$ of degree $2k-1$,

$$\lambda_{2k-1}(x) = \sum_{m=1}^{2k-1} \left(\frac{x}{m}\right) \sum_{p=1}^{k} p! S(k, p) (p+1-p)! S(k, m+1-p)$$

By (1) $\lambda_{2k-1}(n) = \sum_{p=0}^{n} (p^k - p^{k-1}) \left((n+1-p)^k - (n+1-p)^{k-1}\right)$ for all positive integers $n$. We want to have a closer look at this aspect of the polynomials.

**Lemma 7** The functions $\beta_{2k+1}(n) := \sum_{j=1}^{n} j^k (n-j)^k$, $n, k \in \mathbb{N}_0$, can be extended to polynomials $\beta_{2k+1} \in \mathbb{R}[x]$ of degree $2k+1$,

$$\beta_{2k+1}(x) = \frac{x^{2k+1}k!}{(2k+1)!} + (-1)^k \sum_{j=0}^{(k-1)/2} \frac{B_{2k-j}}{k-j} \left(\frac{k}{2j+1}\right) x^{2j+1}$$

where $B(n) = \sum_{j=0}^{n} \sum_{i=0}^{j} \frac{(-1)^i}{j+1} (\frac{j}{i}) i^n$ is the $n$-th Bernoulli number.

**Proof.** The functions $f_{k+1}(n) := \sum_{j=1}^{n} j^k$ can be extended to polynomials $f \in \mathbb{R}[x]$ of degree $k+1$ because $\nabla f_{k+1}(n) = f_{k+1}(n) - f_{k+1}(n-1) = n^k$ for all $n \in \mathbb{N}_1$ can be extended to the polynomial $x^k$. It is well known [2, § 83] that

$$f_{k+1}(n) = \sum_{j=0}^{k} \binom{k+1}{j} B_j \frac{(n+1)^{k+1-j}}{k+1}$$
Hence $\beta_{2k+1}(n+1)$

\[
\begin{align*}
&= \sum_{j=0}^{n} j^k (n+1-j)^k = \sum_{i=0}^{k} \binom{k}{i} (-1)^{k-i} (n+1)^i \sum_{j=1}^{n} j^{2k-i} \\
&= \sum_{i=0}^{k} \binom{k}{i} (-1)^{k-i} (n+1)^i \sum_{j=0}^{2k-i} \binom{2k-i}{j} \frac{B_j (n+1)^{2k-i+1-j}}{2k-i+1} \\
&= \sum_{j=0}^{2k} B_j (n+1)^{2k+1-j} \sum_{i=0}^{2k-j} \binom{k}{i} \binom{2k-i+1}{j} \frac{(-1)^{k-i}}{2k-i+1}.
\end{align*}
\]

Note that for $j = 0$ we obtain the term $B_0 (n+1)^{2k+1} \sum_{i=0}^{k} \binom{k}{i} (-1)^{k-i}$, where

\[
\sum_{i=0}^{k} \binom{k}{i} \frac{(-1)^{k-i}}{2k-i+1} = \sum_{i=0}^{\infty} \binom{k}{i} \frac{(-1)^{i} x^{k+1+i}}{k+1+i} \bigg|_{x=1} = \int_0^1 x^k (1-x)^k \, dx = \text{Beta}(k, k) = \frac{k! k!}{(2k+1)!}.
\]

hence $\beta_{2k+1}(n+1)$

\[
\begin{align*}
&= \frac{(n+1)^{2k+1} k! k!}{(2k+1)!} + \sum_{j=1}^{2k} B_j (n+1)^{2k+1-j} \sum_{i=0}^{2k-j} \binom{k}{i} \binom{2k-i+1}{j} \frac{(-1)^{k-i}}{2k-i+1}.
\end{align*}
\]

For $j > 0$ holds

\[
\begin{align*}
&= \sum_{i=0}^{2k-j} \binom{k}{i} \binom{2k-i+1}{j} \frac{(-1)^{k-i}}{2k-i+1} \\
&= \frac{1}{j} \sum_{i=0}^{2k-j} \binom{k}{i} \left( \binom{2k-i+1}{j} - \binom{2k-i}{j} \right) (-1)^{k-i} \\
&= \frac{(-1)^{k-j-1}}{j} \left( \binom{k-j}{2k+1-j} - \binom{k}{2k-j+1} \right).
\end{align*}
\]

Hence $\beta_{2k+1}(x)$

\[
\begin{align*}
&= \frac{x^{2k+1} k! k!}{(2k+1)!} + \sum_{j=1}^{2k} B_j x^{2k+1-j} \frac{(-1)^{k-j-1}}{j} \left( \binom{k-j}{2k+1-j} - \binom{k}{2k+1-j} \right) \\
&= \frac{x^{2k+1} k! k!}{(2k+1)!} + \sum_{j=1}^{k} B_{2k+1-j} x^j \frac{(-1)^{k-j}}{2k+1-j} \left( \binom{k}{j} (-1)^j - \binom{k}{j} \right)
\end{align*}
\]
Starting at $j = 3$, the Bernoulli numbers $B_j$ are 0 for odd $j$. Therefore, $\beta_{2k+1} (x)$

$$
= x^{2k+1}k!k!\frac{(k-1)!}{(2k+1)!} + \frac{1}{2} \sum_{j=0}^{k-1} B_{2(k-j)}x^{2j+1} \frac{(-1)^{k+1}}{2(k-j)} \left( \frac{2j-k}{2j+1} - \frac{k}{2j+1} \right)
$$

$$
= x^{2k+1}k!k!\frac{(k-1)!}{(2k+1)!} + (-1)^{k+1} \sum_{j=0}^{k} \frac{B_{2(k-j)}}{k-j} \frac{k}{2j+1} x^{2j+1}.
$$

We call the polynomials $\beta_{2k+1} (x)$ beta polynomials, because

$$
\lim_{n \to \infty} \beta_{2k+1} (n) n^{-2k-1} = \frac{k!k!}{(2k+1)!}.
$$

Lemma 7 allows us to improve on this result.

**Corollary 8** If $k$ is odd, then $\beta_{2k+1} (x) / x^k = x^{k+1}k!k!\frac{(k-1)!}{(2k+1)!} - \frac{2B_{k+1}}{k+1} + o(x^k)$. If $k$ is even, then $\beta_{2k+1} (x) / x^{k-1} = x^{k+2}k!k!\frac{(k-1)!}{(2k+1)!} + \frac{2kB_{k+2}}{k+2} + o(x^{k-1})$.

The second order backwards difference connects the beta polynomials to the polynomials $\lambda_{2k-1} (x)$, because

$$
\nabla^2 \beta_{2k+1} (n + 1) = \beta_{2k+1} (n + 1) - 2\beta_{2k+1} (n) + \beta_{2k+1} (n - 1)
$$

$$
= \sum_{s=0}^{n} (s^k - s^{k-1}) \left( (n + 1 - s)^k - (n + 1 - s)^{k-1} \right)
$$

$$
= \lambda_{2k-1} (n)
$$

for all positive integers $n$, and hence

$$
\nabla^2 \beta_{2k+1} (x + 1) = \lambda_{2k-1} (x)
$$

for all $x \in \mathbb{R}$. Therefore we obtain the following asymptotics for $|\Lambda_n^k|$.

**Corollary 9** For odd $k$

$$
|\Lambda_n^k| \approx \frac{((n+1)^{2k+1}-2n^{2k+1}+(n-1)^{2k+1})k!k!}{(2k+1)!} - \frac{2((n+1)^{k+1}-2n^{k+1}+(n-1)^{k+1})B_{k+1}}{k+1}.
$$

For even $k$,

$$
|\Lambda_n^k| \approx \frac{((n+1)^{2k+1}-2n^{2k+1}+(n-1)^{2k+1})k!k!}{(2k+1)!} + \frac{2k((n+1)^{k+1}-2n^{k+1}+(n-1)^{k+1})B_{k+2}}{k+2}.
$$

The number of nonintersecting $k$-tuples of intervals is approximately

$$
\binom{n+1}{2} k - \frac{((n+1)^{2k+1}-2n^{2k+1}+(n-1)^{2k+1})k!k!}{(2k+1)!} + \frac{2((n+1)^{k+1}-n^{k+1})B_{k+1}}{k+1}
$$

for odd $k$, and

$$
\binom{n+1}{2} k - \frac{((n+1)^{2k+1}-2n^{2k+1}+(n-1)^{2k+1})k!k!}{(2k+1)!} - \frac{2((n+1)^{k+1}-n^{k+1})B_{k+2}}{k+2}
$$

for even $k$.

**Proof.** If $k$ is odd, then $\beta_{2k+1} (x) \approx x^{2k+1}k!k!\frac{(k-1)!}{(2k+1)!} - \frac{2x^kB_{k+1}}{k+1}$ according to the previous Corollary, hence

$$
|\Lambda_n^k| = \lambda_{2k-1} (n) \approx
$$
(n+1)^{2k+1} k! k! - 2(n+1)^k B_{k+1} - 2 \left( \frac{n^{2k+1} k! k!}{(2k+1)!} \right) - \frac{2n^k B_{k+1}}{k+1} - \frac{(n-1)^{2k+1} k! k!}{(2k+1)!} - \frac{2(n-1)^k B_{k+1}}{k+1} \right)

Recall that the number of $k$-tuples of subintervals from $[1, n]$ intersecting in $l \geq 1$ points equals the number of $k$-tuples chosen from $[1, n+1-l]$ intersecting in 1 point. Hence the number of nonintersecting $k$-tuples equals \( \binom{n+1}{2}^k - \sum_{l=1}^{n} |\Lambda_{n}^k| \)

\[
\begin{align*}
\binom{n+1}{2} - \sum_{s=1}^{n} s^k (n+1-s)^k + \sum_{s=1}^{n} (s-1)^k (n+1-s)^k \\
\binom{n+1}{2} - \beta_{2k+1} (n+1) + \beta_{2k+1} (n) \\
\approx \binom{n+1}{2} + \frac{(n^{2k+1} - (n+1)^{2k+1}) k! k!}{(2k+1)!} - \frac{2(n^k - (n+1)^k) B_{k+1}}{k+1} 
\end{align*}
\]

For even $k$ the result follows in the same way. 

It is no surprise that the above approximation to $|\Lambda_{n}^k|$ works well, even for small $n$. The following small table shows some relative approximation errors $(|\Lambda_{n}^k| - \text{approximation}) / |\Lambda_{n}^k|$ for $k = 5, 6$, and $n$ between 3 and 6.

\[
\begin{array}{cccccc}
\text{n} & \text{k=5} & \text{k=6} \\
3 & -2.2 \cdot 10^{-3} & -9.7 \cdot 10^{-3} \\
4 & -5.8 \cdot 10^{-4} & 8.0 \cdot 10^{-5} \\
5 & -1.8 \cdot 10^{-4} & -1.0 \cdot 10^{-5} \\
6 & -6.8 \cdot 10^{-5} & 1.8 \cdot 10^{-6} \\
\end{array}
\]

### 5.1 Expansion in Odd Degrees

The beta polynomials are sums of odd powers, and therefore $\lambda_{2k-1}(x) = \nabla^2 \beta_{2k+1}(x+1)$ is a polynomial that contains only odd powers of $x$. The experimental results we mentioned in the introduction let us conjecture that $|\Lambda_{n}^k|$ can be expanded in terms of binomial coefficients of odd degrees,

\[
|\Lambda_{n}^k| = \sum_{j=1}^{k} \binom{n+j-1}{2j-1} d(j, k)
\]

for some positive integers $d(j, k)$. We will now determine those coefficients in terms of the numbers $c(m, k)$ we discussed earlier (see (??)). First we need one more property of $c(m, k)$.

**Lemma 10** For all positive integers $m$ holds

\[
c(m, k) = \sum_{i=m}^{2k-1} (-1)^{i-1} c(i, k) \binom{i-1}{m-1}
\]
Proof. We noted that the polynomials \( \lambda_{2k-1}(x) \) are odd, \( \lambda_{2k-1}(x) = -\lambda_{2k-1}(-x) \). By expansion (4) \( \lambda_{2k-1}(n) \)

\[
\begin{align*}
&= \sum_{m=1}^{2k-1} \binom{n}{m} c(m, k) = -\lambda_{2k-1}(-n) = -\sum_{m=1}^{2k-1} \left( \frac{-n}{m} \right) c(m, k) \\
&= \sum_{i=1}^{2k-1} (-1)^{i-1} c(i, k) \binom{n + i - 1}{i} \\
&= \sum_{i=1}^{2k-1} (-1)^{i-1} c(i, k) \sum_{m=1}^{2k-1} \binom{n}{m} \binom{i - 1}{i - m} \\
&= \sum_{m=1}^{2k-1} \binom{n}{m} \sum_{i=m}^{2k-1} (-1)^{i-1} c(i, k) \binom{i - 1}{m - i}
\end{align*}
\]

Comparing coefficients of the polynomial basis \( \binom{n}{m} \) shows that

\[
c(m, k) = \sum_{i=m}^{2k-1} (-1)^{i-1} c(i, k) \binom{i - 1}{m - 1}.
\]

Lemma 11 Let \( d(j, k) := \sum_{i=1}^{2k-1} (-1)^{i-1} \frac{j-i}{j-i-1} c(i, k) \) for all \( j = 1, \ldots, k \). Then

\[
|\Lambda_n^k| = \sum_{j=1}^{k} \binom{n + j - 1}{2j - 1} d(j, k)
\]

Note that \( \frac{i-j}{i-j-1} \) must be interpreted as 1 if \( j = 1 \), and as \( \frac{i-j-1}{j-1} \) if \( i = j > 1 \).

Proof. Because of

\[
\begin{align*}
\sum_{j \geq 1} \binom{n + j - 1}{2j - 1} d(j, k) &= \sum_{j \geq 1} \sum_{m=0}^{2j-1} \binom{n}{m} \binom{j-1}{2j-1-m} d(j, k) \\
&= \sum_{m=0}^{2k-1} \binom{n}{m} \sum_{j=1}^{m} \binom{j-1}{m-j} d(j, k)
\end{align*}
\]

and \( |\Lambda_n^k| = \sum_{m=1}^{2k-1} \binom{n}{m} c(m, k) \) it suffices to show that

\[
c(m, k) = \sum_{j=1}^{m} \binom{j-1}{m-j} d(j, k) \\
= \sum_{j=1}^{m} \binom{j-1}{m-j} \sum_{i=1}^{2k-1} (-1)^{i-1} \frac{i-1}{i-j} \binom{i-j}{j-1} c(i, k)
\]
Combine the identity $\sum_{j=1}^{m} \frac{j-1}{m-j} = \binom{i-1}{m-1}$ and the previous Lemma to get
\[
c(m, k) = \sum_{i=1}^{2k-1} (-1)^{i-1} c(i, k) \sum_{j=1}^{m} \frac{j-1}{m-j} \binom{i-j}{j-1} = \sum_{j=1}^{m} \frac{j-1}{m-j} \sum_{i=1}^{k} (-1)^{i-1} \binom{i-j}{j-1} c(i, k)
\]
as desired. □

6 Distribution of the Size of the Intersection

If we consider the cardinality of the intersection of a sequence of $k$ random (equally likely) subintervals of $[1, n]$ as a random variable $L_n$, say, then
\[
\Pr(L_n = l) = \Pr(L_{n+1-l} = 1) = \left| \Lambda_{n+1-l} \right| \binom{n+1}{2}^{-k}
\]
for all positive integers $l$. For the graph below we applied Corollary 9, plotting a continuous approximation to this distribution ($n = 100$ and $k = 5$). In that example, $\Pr(L_n = 0) = 0.876$.

Pr $(L_{100} = l)$ for $1 \leq l \leq 50$ and $k = 5$ (continuous approximation)

The probability of no intersection will increase with the number $k$ of intervals; by 2 and Corollary 9
\[
\Pr(L_n = 0) = 1 - \left( \frac{n+1}{2} \right)^{-k} \sum_{p=1}^{n} \left( p^k - (p-1)^k \right) (n+1-p)^k \\
\approx 1 - 2^k \frac{(n+1)^{2k+1} - n^{2k+1}}{(2k+1)!n^k (n+1)^k} \approx 1 - \frac{k!k!}{(2k)!} \\
\approx 1 - 2^k \frac{k^{2k+1} \sqrt{2\pi}}{(2k)^{2k+1/2}} \approx 1 - 2^{-k} \sqrt{k\pi}
\]
(10)
The probability of selecting $k$ nonintersecting intervals

### 6.1 The First Two Moments

Calculating the expected value $\mu_{k,n}$ of $L_n$ is a straightforward exercise; finding the variance $\sigma^2_{k,n}$ may be easier via the second (falling) factorial moment (Lemma 13).

**Lemma 12** The expected size $\mu_{k,n}$ of the intersection of a $k$-tupel of subintervals of $[1, n]$ equals

$$\left( \frac{n + 1}{2} \right)^{-k} \sum_{i=1}^{n} i^k (n + 1 - i)^k.$$  

Furthermore, $\lim_{n \to \infty} \mu_{k,n} \approx \frac{2^{k+1} n^{k+1} k!}{n^{k+1} 2^{k} (2k+1)!}$

$$\approx (n + 1)^{k+1} n^{-k} 2^{-k-1} \sqrt{\pi / k}.$$  

With the help of Stirling’s formula we can further approximate $\mu_{k,n}$ by $(n + 1)^{k+1} n^{-k} 2^{-k-1} \sqrt{\pi / k}$. There is however a significant loss in precision from the first to the second approximation. For example, if $k = n = 50$, the relative error $(\mu_{k,n} - (n + 1)^{k+1} n^{-k} 2^{-k-1} \sqrt{\pi / k}) / \mu_{k,n}$ equals $-2 \times 10^{-32}$, and $(\mu_{k,n} - (n + 1)^{k+1} n^{-k} 2^{-k-1} \sqrt{\pi / k}) / \mu_{k,n} = -7 \times 10^{-3}$. In any case, we observe a slow increase of the expected cardinality of the intersection when $n$ increases, and an exponential decline, when $k$ increases.

**Proof.** From $\Pr (L_n = l) = \Pr (L_{n+1-l} = 1) = |\Lambda_{n+1-l}^k (\begin{array}{c} n+1 \end{array})^{-k} |$ for all positive integers $l$ follows $\mu_{k,n} = \left( \frac{n + 1}{2} \right)^{-k} \sum_{l=1}^{n} |\Lambda_{n+1-l}^k |$. In terms of beta polynomials we must show that

$$\beta_{2k+1} (n+1) = \sum_{l=1}^{n} l \nabla^2 \beta_{2k+1} (n + 2 - l)$$  

(11)
(see (9)). This holds for \( n = 1 \) because \( \nabla^2 \beta_{2k+1} (2) = |A_k| = 1 = 1^k (2-1)^k = \beta_{2k+1} (2) \).

Note that for \( k > 0 \) the beta polynomials have roots at \( n = 0 \) and \( n = 1 \). By induction

\[
\sum_{l=1}^{n+1} l \nabla^2 \beta_{2k+1} (n + 2 - l)
= \sum_{l=0}^{n} (l + 1) \nabla^2 \beta_{2k+1} (n + 1 - l) = \beta_{2k+1} (n) + \nabla^2 \sum_{l=0}^{n-1} \beta_{2k+1} (n + 1 - l)
= \beta_{2k+1} (n) + \nabla (\beta_{2k+1} (n + 1) - \beta_{2k+1} (1))
= \beta_{2k+1} (n) + \beta_{2k+1} (n + 1) - \beta_{2k+1} (n)
\]

We saw in Lemma 7 that \( \beta_{2k+1} (n + 1) \approx \frac{(n+1)^k}{k! (2k+1)!} \), explaining the approximation \( \frac{2^k (n+1)^{k+1} k l k!}{n^{k+1} (2k+1)!} \)
for \( \mu_{k,n} \).

**Lemma 13**

\[
E [L_n (L_n - 1)] = 2 \left( \begin{array}{c} n+1 \\ 2 \end{array} \right)^{-k} \sum_{i=1}^{n} \sum_{j=1}^{n-i} (ij)^k
\]

**Proof.** Note that \( L_1 \equiv 1 \), thus the Lemma holds for \( n = 1 \), because \( 0 = E [L_1 (L_1 - 1)] = 2^{(2)} \left( \begin{array}{c} n+1 \\ 2 \end{array} \right)^{-k} \sum_{i=1}^{1} \sum_{j=1}^{1-i} (ij)^k \). Suppose it holds for \( n \geq 1 \). By induction,

\[
\left( \begin{array}{c} n+2 \\ 2 \end{array} \right)^{-k} E [L_{n+1} (L_{n+1} - 1)]
= \sum_{l=1}^{n+1} l (l - 1) |A_{n+2-l}^k| = \sum_{l=1}^{n} (l - 1) l |A_{n+1-l}^k| + \sum_{l=1}^{n} 2l |A_{n+1-l}^k|
= \left( \begin{array}{c} n+1 \\ 2 \end{array} \right)^{k} E [L_n (L_n - 1)] + 2 \beta_{2k+1} (n + 1)
= 2 \sum_{i=0}^{n} i^k \sum_{j=i}^{n} (j-i)^k + 2 \sum_{i=0}^{n+1} i^k (n+1-i)^k = 2 \sum_{i=0}^{n+1} i^k \sum_{j=i}^{n+1} (j-i)^k
\]


\[
= \left( \begin{array}{c} n+2 \\ 2 \end{array} \right)^{-k} \left( 2 \sum_{i=1}^{n} \sum_{j=1}^{n-i} (ij)^k - \beta_{2k+1} (n + 1) (\beta_{2k+1} (n + 1) - 1) \right).
\]

**References**


