Sampling of Infinite Set by J-sequence
: Solving Bertrand's (Chord) Paradox

Jeonghoon LEE
Seoul, Korea
oldfuture11@naver.com
July 12, 2022

Abstract :

In an infinite set, probabilities are defined on the structure of the set rather than on individual elements. We should take into account the property of a σ-algebra where probabilities are defined. A σ-algebra is closed under 'only countable' unions, and the axioms of probability assume σ-additivity. If this is overlooked, something bizarre could be happened as the proposed three solutions of Bertrand's problem.

Bertrand's problem is not a paradox, but well defined(posed).

The suggested three solutions have the common problem of dividing the sample space into an uncountably infinite number of sets and treating them equally. If a set is divided into equal(treated) and uncountable infinity, all the divided sets have probability 0, so calculating conditional probabilities with these sets or comparing them with each other becomes meaningless.

In a sample space composed of an uncountably infinite number of elements such as [0,1], after calculating the number of cases using the sets(J-sequence m-collection cover) generated by equally dividing the sample space into finite numbers, the probability of an event can be calculated with its limit value(as m becomes infinite). The answer of Bertrand's problem is 1/3.

1. Defining Probability of an Infinite Set

Infinite sets have characteristics that allow one-to-one correspondence with their proper subset. For example, the set of natural numbers can correspond 1-1 to the set of even numbers in such a way that x corresponds to $2x$. In the set of natural numbers, probabilities can not be defined if each natural number has the same probability.

If the probability of one natural number is P, for every natural number, the probability should be defined as P. In this case, if P=0, the probability of the set of natural numbers as sample space becomes 0, and if it is greater than 0, the probability of the sample space becomes infinite, which goes against the axiom of probability that the probability of the sample space should be 1.

Then, how is the probability defined in an infinite set? In an infinite set, probabilities are defined on the structure(e.g. topology) of the set rather than on individual elements.

When there is a structure in a set and σ-algebra is defined from the structure, a basis for defining probabilities is established.

The definition of σ-algebra is as follows.

1.1. Definition of σ-algebra ([4]Rudin 1987)

Let X be some set, and let P(X) represent its power set.

Then a subset $\Sigma \subseteq P(X)$ *is called a σ-algebra if it satisfies the following three properties:*

1.11. X is in Σ, and X is considered to be the universal set in the following context.

1.12. Σ is closed under complementation: If A is in Σ, then so is its complement, $X \setminus A$.

1.13. Σ is closed under countable unions: If A1, A2, A3, ... are in Σ, then so is $A = A1 \cup A2 \cup A3 \cup \cdots$.

1.2. It should be noted here that σ-algebra is closed under 'only countable' unions and the axiom of probability assume σ-additivity.

Strange things will be happened, if σ-algebra is closed under uncountable unions of sets in which the probability is defined. For exemple, if we could define a probability of $\bigcup_{\alpha \in S} A_\alpha$ (where S is an uncountable set, and A_α are pairwise disjoint) and the probability of each A_α is greater than 0, then, the probability of $\bigcup_{\alpha \in S} A_\alpha$ should be infinite.

2. The Problem of Sampling of Infinite Sets

As discussed above, the structure of σ-algebra and the property of σ-additivity are necessary to define a probability of an infinite set. If this is overlooked when sampling to calculate the probability, something bizarre could be happened. The problem called Bertrand's paradox is an example of this phenomenon.

2.1. Bertrand's paradox

It was proposed by Bertrand, Joseph in his book 'Calcul des probabilités (1889)'

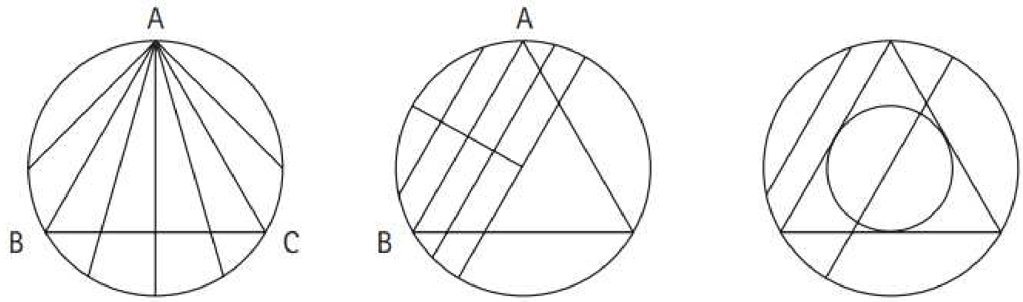
Bertrand's Problem: Consider a circle with an equilateral triangle inscribed in it. What is the probability that a chord selected at random will be longer than the side of the triangle? ([2]Drory 2015)

2.11. The suggested three solutions ([6]Clark 2012)

(solution 1) *The chords from a vertex of the triangle to the circumference are longer if they lie within the angle at the vertex. Since that is true of one-third of the chords, the probability is one-third.*

(solution 2) *The chords parallel to one side of such a triangle are longer if they intersect the inner half of the radius perpendicular to them, so that their midpoint falls within the triangle. So the probability is one-half.*

(solution 3) *A chord is also longer if its midpoint falls within a circle inscribed within the triangle. The inner circle will have a radius one-half and therefore an area one-quarter that of the outer one. So the probability is one-quarter.* ([6]Clark 2012:22)



([6]Clark 2012:23)

2.2. Problems of the three proposed solutions

2.21. Solution 1

A chord is a line segment connecting two points on the circumference. Therefore, {a}×{b}∈S×S represents one chord, where a∈ S, b∈S, S is a circle.

The logic of Solution 1 is as follows.

Solution 1 divides the sample space equally into the uncountably infinite number of the sets.

Let E_α=S×{α} be the divided set by one point α of the sample space X=S×S, where S is a circle, then $X=\bigcup_{\alpha\in S} E_\alpha$ and E_α are pairwise disjoint.

Now for the event B ⊂ S×S that satisfies the condition, then

B = X∩B =$(\bigcup_{\alpha\in S} E_\alpha)$∩B = $\bigcup_{\alpha\in S}(E_\alpha$∩B).

Since the conditional probability P(B | E_α) = 1/3 according to the solution, and P(E_α∩B) = P(B | E_α)*P(E_α), thus,

P(B)=P($\bigcup_{\alpha \in S}$(E_α∩B))=$\sum_{\alpha \in S}$P(E_α∩B)= $\sum_{\alpha \in S}$P(B | E_α)*P(E_α)= 1/3$\sum_{\alpha \in S}$P(E_α) =1/3.

This is the logic of Solution 1.

However, since $\bigcup_{\alpha \in S}$(E_α∩B) is an uncountable unions of sets, P($\bigcup_{\alpha \in S}$(E_α∩B)) cannot be equal to $\sum_{\alpha \in S}$P(E_α∩B). Actually, $\sum_{\alpha \in S}$P(E_α∩B) is not well defined.

This is because, as discussed earlier, the probability is not well defined in the uncountable unions of sets, where the probability of each set is not zero. Therefore, it should be P(E_α∩B)=0, and the above logic has no foundation. (∵If each P(E_α∩B) > 0, then P($\bigcup_{\alpha \in S}$(E_α∩B)) becomes infinite)

2.22. Solution 2

The entire sample space is partitioned based on the distance between the center of the circle and the midpoint of the chord connecting two points, and the probability is calculated based on the length.

The sample space X = S×S, where S is a circle of radius $\frac{1}{2\pi}$, is partitioned according to distance ‘α’ between the center of the circle and the midpoint of the chord connecting {a} and {b}, where {a}×{b}∈X. If E_α is a partition of distance ‘α’, then X = $\bigcup_{0 \le \alpha \le 1/2\pi} E_\alpha$.

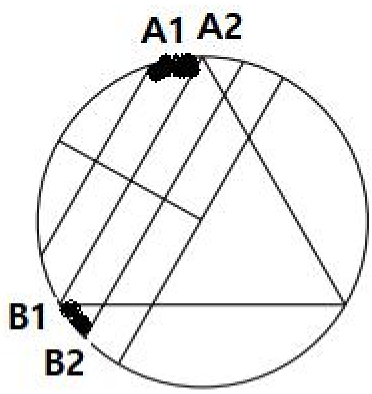
In this case, it is divided into two sets based on α. That is, if X_1= $\bigcup_{0 \le \alpha \le 1/4\pi} E_\alpha$, X_2= $\bigcup_{1/4\pi < \alpha \le 1/2\pi} E_\alpha$, then X= $\bigcup_{0 \le \alpha \le 1/2\pi} E_\alpha$ = $X_1 \cup X_2$.

In this case, the event B ⊂ S×S satisfying the condition becomes B=X_1. Since P(B) is the ratio of the radius length assigned to the set B to the total radius length, thus P(B) = 1/2. More precisely, where B=X_1= $\bigcup_{0 \le \alpha \le 1/4\pi} E_\alpha$ and, P(B) =P($\bigcup_{0 \le \alpha \le 1/4\pi} E_\alpha$), applying the indifference principle here, P(E_α) is constant regardless of α, so it is proportional to the assigned length. This is the logic of Solution 2.

However, as discussed earlier, 0≤α≤1/4π is an uncountable infinity, so it cannot be P($\bigcup_{0 \le \alpha \le 1/4\pi} E_\alpha$)= $\sum_{0 \le \alpha \le 1/4\pi} P(E_\alpha)$.

It should be P(E_α)=0, therefore, the logic that the probability is proportional to the assigned length has no foundation.

If you draw a picture, this solution does not really conform to the principle of indifference. If the same logic is applied by dividing the radius of a circle based on a small line segment instead of a point, the corresponding circumference length is small when it is close to the center of the circle for a line segment of the same size, but when it is far from the center of the circle, the corresponding circumference length is get bigger.



Looking at the length of the circumference corresponding to the same radius length in the figure, it can be seen that A1A2 (farther from the center of the circle) is much longer than B1B2 (close to the center of the circle). However, Solution 2 treats the probability of selecting a point from B1B2 as equal to the probability of selecting one from A1A2. The second solution overlooked this problem by thinking in terms of points.

2.23. Solution 3

The entire sample space is divided based on the midpoint of the chord, and the probability is calculated with the area of the set of midpoints.

The sample space X = S×S, where S is a circle of radius $\frac{1}{2\pi}$, is partitioned according to ‘α’, the midpoint of the chord connecting {a} and {b}, where {a}×{b}∈X. If E_α is a partition of ‘α’, then $X=\bigcup_{\alpha\in A} E_\alpha$, where A is the interior of S.

Now, if the interior of the inner circle (radius 1/2π) is C, and the remainder is D, where $X_1=\bigcup_{\alpha\in C} E_\alpha$, $X_2=\bigcup_{\alpha\in D} E_\alpha$, then $X = X_1 \cup X_2$.

In this case, the event B ⊂ S×S that satisfies the condition becomes B=X_1, P(B) is the ratio of the area of C to the area of the whole circle, and since the radius of C is 1/2 of the whole circle, the area is 1/4 of the whole circle, thus P(B) = 1/4. More precisely, where B=$X_1=\bigcup_{\alpha\in C} E_\alpha$, and

P(B) = P($\bigcup_{\alpha\in C} E_\alpha$), applying the indifference principle here, P(E_α) is constant regardless of α, so P(B) is proportional to the area of C. That is the logic of the solution.

However, as discussed earlier, since C consists of an uncountably infinite number of points, it cannot be P($\bigcup_{\alpha\in C} E_\alpha$) = $\sum_{\alpha\in C}$P(E_α).

It should be P(E_α)=0, therefore, the logic that the probability is proportional to the area of C has no foundation.

If you draw a picture, this solution does not really conform to the principle of indifference. If the same logic is applied by dividing the inside of the circle S based on a small 'ball' rather than a 'point', for a ball of the same size, when this ball is at the center of the circle S, the length of the corresponding circumference is the same in any direction but, we see that as the ball moves

away from the center of the circle S, the length of the corresponding circumference varies with the direction, and the corresponding circumference is always the same or longer than when the ball is at the center of the circle.
It treats different probabilities as equal. Since this solution is point-based, it overlooks this difference.

2.3. The suggested three solutions have the common problem of dividing the sample space into uncountably infinite number of sets and treating them equally.

If a set is divided into equal and uncountable infinity, all the divided sets have probability 0, so calculating conditional probabilities with these sets or comparing them with each other becomes meaningless.
That is, without considering σ-algebra and σ-additivity, the error of calculating the probability in a set with zero probability and adding probabilities of the sets by an uncountably infinite number of times.

In order to solve this problem, we will introduce a new concept called 'J-sequence' for correct sampling considering the structure of σ-algebra and σ-additivity, where the probability is defined.

3. J-sequence : the number of cases using J-sequence m-collection cover

3.1. The Axioms of Probability

The followings are generally accepted as the axioms of probability :

3.11. The probability of an event(E) is non negative, that is, P(E)≥0

3.12. Probability of the entire sample space(U) is 1, that is P(U)=1

3.13. The assumption of σ-additivity:

E_i are disjoint events, then $P(\bigcup_{i=1}^{\infty} E_i) = \sum_{i=1}^{\infty} P(E_i)$

If we consider [0,1]⊂R (real number) as sample space, Lebesgue measurable set as an event and the Lebesgue measure as probability, it is consistent with the probability axioms.

3.2. Definition: J-sequence

J-sequence $\{a_k\}$ is defined as follows :

$a_1=0$, $a_2=1/2$

$a_{2^n+i}= a_i +1/2^{n+1}$ $(1\leq i\leq 2^n,\ n\geq 1)$

If you list J-sequence, that is

0, 1/2, 1/4, 3/4, 1/8, 3/8, 5/8, 7/8, 1/16, 3/16, 5/16, 7/16, 9/16, 11/16, 13/16, 15/16,....

3.3. Definition: J-sequence m-collection

J-Sequence m-collection S_{2^m}, is defined by

$S_{2^m}=\{a_1, a_2, a_3, \ldots , a_{2^m}\}=\{0/2^m, 1/2^m, 2/2^m,\ldots ,(2^m-1)/2^m\}$,

where $\{a_k\}$ is J-sequence.

3.4. Definition: J-sequence m-collection cover

J-sequence m-collection cover J_{2^m} is defined by

$$J_{2^m}=\{I_i | I_i=(\frac{i-1}{2^m}, \frac{i}{2^m})\subset[0,1],\ 1\leq i\leq 2^m\}$$

3.5. Definition: J-sequence m-collection cover counting about set A

J-sequence m-collection cover counting about set A is defined by n(Q_{2^m}(A)) (the number of elements of Q_{2^m}(A)), where J_{2^m} is J-sequence m-collection cover and $Q_{2^m}(A)= \{I_i \mid I_i\in J_{2^m},\ \mu(I_i\text{-}A)=0\}$

($\mu(I_i$-A) is the Lebesgue measure of (I_i-A)=($I_i\cap A^c$))

※ From now on, μ(X) denotes the Lebesgue measure of X.

3.6. Proposition 1

For every open interval A= (a,b)$\subseteq$[0,1],

$$\lim_{2^m\to\infty} n(Q_{2^m}(A))/2^m = b-a = \mu(A)$$

(proof)

For every open interval A= (a,b) $\subseteq$ [0,1],

$\exists$ s, t $\in N$ (natural number) such that

$s/2^m \le a < (s+1)/2^m$, $t/2^m \le b < (t+1)/2^m$,

then, $(t-s-1)/2^m \le (b-a) \le (t-s+1)/2^m$

Since $I_i \subset (a,b)=A$ for $(s+2) \le i \le (t-1)$, then $I_i = (\frac{i-1}{2^m}, \frac{i}{2^m}) \in Q_{2^m}(A)$.

(where $Q_{2^m}(A) = \{ I_i \mid I_i \in J_{2^m},\ \mu(I_i - A) = 0 \}$)

If $i \le s$ or $i \ge (t+2)$,

then $I_i \cap A = \varnothing$, thus $\mu(I_i - A) = \mu(I_i) = 1/2^m$, so $I_i \notin Q_{2^m}$

$\therefore$ $(t-s-2) \le n(Q_{2^m}(A)) \le (t-s+1)$

$\Leftrightarrow$ $(t-s-2)/2^m \le n(Q_{2^m}(A))/2^m \le (t-s+1)/2^m$

$\rightarrow$ $\lim_{2^m \to \infty} (t-s-2)/2^m \le \lim_{2^m \to \infty} n(Q_{2^m}(A))/2^m \le \lim_{2^m \to \infty} (t-s+1)/2^m$

$\rightarrow$ $(b-a) \le \lim_{2^m \to \infty} n(Q_{2^m}(A))/2^m \le (b-a)$ ($\because \lim_{2^m \to \infty} (t-s)/2^m = (b-a)$)

$\therefore \lim_{2^m \to \infty} n(Q_{2^m}(A))/2^m = b-a = \mu(A)$

3.7. Lemma 1

For every B= (a,b) $\subseteq$ [0,1] and for any $m \in N$ (natural number),

$n(Q_{2^m}(B))/2^m \le \mu(B)$.

(proof)

Where $J_{2^m} = \{ I_i | I_i = (\frac{i-1}{2^m}, \frac{i}{2^m}) \subset [0,1],\ 1 \le i \le 2^m \}$ and

$Q_{2^m}(B) = \{ I_i \mid I_i \in J_{2^m},\ \mu(I_i - B) = 0 \}$,

since $\mu(I_i) = 1/2^m$, and $I_i (1 \le i \le 2^m)$ are pairwise disjoint,

$$n(Q_{2^m}(A))/2^m = \sum_{I_i \in Q_{2^m}(B)} \mu(I_i) = \mu(\bigcup_{I_i \in Q_{2^m}(B)} I_i) \le \mu(B)$$

$$(\because \bigcup_{I_i \in Q_{2^m}(B)} \subset B)$$

3.8. Lemma 2

If A∩B=∅, then n(Q_{2^m}(A∪B))/2^m ≥ n(Q_{2^m}(A))/2^m +n(Q_{2^m}(B))/2^m

(proof)

For any I_i, since A∩B=∅, there cannot be μ(I_i-A)=0 and μ(I_i-B)=0 at the same time. (where $I_i \in J_{2^m}$). That is, I_i is never counted twice as n(Q_{2^m}(A)) and n(Q_{2^m}(B)).

On the other hand, I_i counted in n(Q_{2^m}(A)) or in n(Q_{2^m}(B)) is necessarily counted in Q_{2^m}(A∪B).

That is, n(Q_{2^m}(A∪B)) ≥ n(Q_{2^m}(A))+ n(Q_{2^m}(B))

∴ n(Q_{2^m}(A∪B))/2^m ≥ n(Q_{2^m}(A))/2^m +n(Q_{2^m}(B))/2^m

3.9. Theorem 1

For every open set B ⊆[0,1], $\lim_{2^m \to \infty}$ n(Q_{2^m}(B))/2^m = μ(B)

(proof)

Every open set B ⊆[0,1] is a countable, disjoint union of open intervals. That is B = $\bigcup_{k=1}^{\infty} A_k$, ($A_k$=($a_k$, b_k)⊆[0,1], and if i≠j, then $A_i \cap A_j$= ∅)

By Lemma 1,

n(Q_{2^m}(B))/2^m ≤ μ(B)

∴ $\lim_{2^m \to \infty}$ n(Q_{2^m}(B))/2^m ≤ μ(B)

On the other hand, since $\sum_{k=1}^{\infty} \mu(A_k)$ = μ(B) , for ∀ε > 0, ∃N such that for all q ≥ N, μ(B)-ε ≤ $\sum_{k=1}^{q} \mu(A_k)$.

Futhermore, by Lemma 2, $\sum_{k=1}^{q}$n($Q_{2^m}(A_k)$)/2^m ≤ n($Q_{2^m}(\bigcup_{k=1}^{q} A_k)$)/$2^m$, by Proposition 1, $\mu(A_k) = \lim_{2^m \to \infty} \mathrm{n}(Q_{2^m}(A_k))/2^m$

$$\therefore \sum_{k=1}^{q} \mu(A_k) = \sum_{k=1}^{q} \lim_{2^m \to \infty} \mathrm{n}(Q_{2^m}(A_k))/2^m \text{ (by Proposition 1)}$$

$$= \lim_{2^m \to \infty} \sum_{k=1}^{q} \mathrm{n}(Q_{2^m}(A_k))/2^m$$

$$\leq \lim_{2^m \to \infty} \mathrm{n}(Q_{2^m}(\bigcup_{k=1}^{q} A_k))/2^m \text{ (by Lemma 2)}$$

$$\leq \lim_{2^m \to \infty} \mathrm{n}(Q_{2^m}(\mathrm{B}))/2^m \quad (\because \bigcup_{k=1}^{q} A_k \subset \mathrm{B})$$

$$\therefore \mu(\mathrm{B}) - \varepsilon \leq \sum_{k=1}^{q} \mu(A_k) \leq \lim_{2^m \to \infty} \mathrm{n}(Q_{2^m}(\mathrm{B}))/2^m$$

From the above, $\mu(\mathrm{B}) - \varepsilon \leq \lim_{2^m \to \infty} \mathrm{n}(Q_{2^m}(\mathrm{B}))/2^m \leq \mu(\mathrm{B})$,

$\varepsilon > 0$ is arbitrary, so $\lim_{2^m \to \infty} \mathrm{n}(Q_{2^m}(\mathrm{B}))/2^m = \mu(\mathrm{B})$.

3.10. Corollary

If $\mathrm{C} = (\mathrm{B} \cup \mathrm{E}) - \mathrm{F}$ (where B is open, $\mu(\mathrm{E}) = \mu(\mathrm{F}) = 0$, and C, B, $\mathrm{F} \subseteq [0,1]$),

then, $\lim_{2^m \to \infty} \mathrm{n}(Q_{2^m}(\mathrm{C}))/2^m = \mu(\mathrm{C})$.

(proof) For every I_i, $\mu(I_i - \mathrm{C}) = 0$ if and only if $\mu(I_i - \mathrm{B}) = 0$.

3.11. Intuition

The method of J-sequence and other concepts, is essentially equivalent to calculating the supremum of measures of all open sets included in the probability set. However, for a measurable set, the Lebesgue measure is the same as the infimum of measures of all open sets containing it. Therefore, there could be some measurable sets to which the method cannot be applied.

4. J-sequence of 2 dimensional measure

4.1. Definition: J-sequence m-collection of 2 dimension

J-sequence m-collection cover of 2 dimension $S_{2^m}\times S_{2^m}$ is defined by $S_{2^m}\times S_{2^m}=\{(a_i,a_j)\in[0,1]\times[0,1] |\ \ a_i\in S_{2^m},\ \ a_j\in S_{2^m}\}$ (where S_{2^m} is J-sequence m-collection)

4.2. Definition: J-sequence m-collection cover of 2 dimension

J-sequence m-collection cover of 2 dimension $J_{2^m}\times J_{2^m}$ is defined by $J_{2^m}\times J_{2^m}=\{\ \ I_i\times I_j\subset[0,1]\times[0,1]\ \mid\ I_i\in J_{2^m},\ \ I_j\in J_{2^m}\}$, where J_{2^m} is J-sequence m-collection cover.

4.3. Definition:

J-sequence m-collection cover counting of 2 dimension about set A.

J-sequence m-collection cover counting of 2 dimension about set A, is defined by n($Q_{2^m\times 2^m}$(A)) (the number of elements of $Q_{2^m\times 2^m}$(A)), where $J_{2^m}\times J_{2^m}$ is J-sequence m-collection cover of 2 dimension and $Q_{2^m\times 2^m}$(A) = $\{I_i\times I_j \mid I_i\times I_j \in J_{2^m}\times J_{2^m},\ \mu(I_i\times I_j\text{-A})=0\}$.

4.4. Proposition 2

For every open rectangle A= (a,b)×(c,d) ⊆ [0,1]×[0,1],

$$\lim_{2^m\to\infty} n(Q_{2^m\times 2^m}(A))/2^{2m} = (b-a)(d-c) = \mu(A).$$

(proof)

For any open rectangle

A= (a,b)×(c,d) ⊆ [0,1]×[0,1],

∃ s,t,u,v ∈ N (natural number) such that

s/2^m ≤ a < (s+1)/2^m , t/2^m ≤ b < (t+1)/2^m,

u/2^m ≤ c < (u+1)/2^m , v/2^m ≤ d < (v+1)/2^m,

then,

$(t\text{-}s\text{-}1)/2^m \le (b\text{-}a) \le (t\text{-}s+1)/2^m$

$(v\text{-}u\text{-}1)/2^m \le (d\text{-}c) \le (v\text{-}u+1)/2^m$

Since $I_i \times I_j \subset (a,b)\times(c,d) = A$, for $(s+2) \le i \le (t-1)$, $(u+2) \le j \le (v-1)$, then $I_i \times I_j \in Q_{2^m \times 2^m}(A)$ (where $Q_{2^m \times 2^m}(A) = \{I_i \times I_j \mid I_i \times I_j \in J_{2^m} \times J_{2^m},\ \mu(I_i \times I_j - A) = 0\}$)

Futhermore,
if $i \le s$ or $i \ge (t+2)$ or $j \le u$ or $j \ge (v+2)$,
then $(I_i \times I_j) \cap A = \varnothing$, thus $\mu(I_i \times I_j - A) = \mu(I_i \times I_j) = 1/2^{2m}$,
so $I_i \times I_j \notin Q_{2^m \times 2^m}(A)$.

$\therefore \quad (t\text{-}s\text{-}2)(v\text{-}u\text{-}2) \le n(Q_{2^m \times 2^m}(A)) \le (t\text{-}s+1)(v\text{-}u+1)$

$\Leftrightarrow \quad (t\text{-}s\text{-}2)(v\text{-}u\text{-}2)/2^{2m} \le n(Q_{2^m \times 2^m}(A))/2^{2m} \le (t\text{-}s+1)(v\text{-}u+1)/2^{2m}$

$\rightarrow \quad \lim_{2^m \to \infty} (t\text{-}s\text{-}2)(v\text{-}u\text{-}2)/2^{2m} \le \lim_{2^m \to \infty} n(Q_{2^m \times 2^m}(A))/2^{2m} \le \lim_{2^m \to \infty} (t\text{-}s+1)(v\text{-}u+1)/2^{2m}$

$\rightarrow \quad (b\text{-}a)(d\text{-}c) \le \lim_{2^m \to \infty} n(Q_{2^m \times 2^m}(A))/2^{2m} \le (b\text{-}a)(d\text{-}c)$

$(\because \lim_{2^m \to \infty} (t\text{-}s)/2^m = (b\text{-}a),\ \lim_{2^m \to \infty} (v\text{-}u)/2^m = (d\text{-}c)\ )$

$\therefore \quad \lim_{2^m \to \infty} n(Q_{2^m \times 2^m}(A))/2^{2m} = (b\text{-}a)(d\text{-}c) = \mu(A)$

4.5. Theorem 2

For every open set $B \subseteq [0,1]\times[0,1]$,

$\lim_{2^m \to \infty} n(Q_{2^m \times 2^m}(B))/2^{2m} = \mu(B)$

(proof)

Every open set $B \subseteq [0.1]\times[0,1]$ is countable unions of almost disjoint closed rectangles(where ‘almost disjoint’ means their interiors are disjoint)

That is, $B = \bigcup_{k=1}^{\infty} \overline{R_k}$, ($\overline{R_k} = [r_{k1}, r_{k2}]\times[r_{k3}, r_{k4}] \subseteq [0.1]\times[0,1]$), and

if $R_k = (r_{k1}, r_{k2})\times(r_{k3}, r_{k4})$, then $\overline{R_k} = R_k \cup \partial \overline{R_k}$ and $\mu(\partial R_k) = 0$

(where $\partial\overline{R_k} = \overline{R_k} - R_k$), and R_k are pairwise disjoint,

so $\mu(B) = \mu(\bigcup_{k=1}^{\infty}\overline{R_k}) = \mu(\bigcup_{k=1}^{\infty}R_k) = \sum_{k=1}^{\infty}\mu(R_k)$.

By the same method of Lemma 1,

$n(Q_{2^m\times 2^m}(B))/2^{2m} \leq \mu(B)$

On the other hand, since $\sum_{k=1}^{\infty}\mu(R_k) = \mu(B)$, for $\forall\varepsilon > 0$, $\exists N$ such that for all $q \geq N$, $\mu(B)-\varepsilon \leq \sum_{k=1}^{q}\mu(R_k)$.

Furthermore, by the same method of Lemma 2,

$\sum_{k=1}^{q} n(Q_{2^m\times 2^m}(R_k))/2^{2m} \leq n(Q_{2^m\times 2^m}(\bigcup_{k=1}^{q})R_k)/2^{2m}$,

and by Proposition 2, $\lim_{2^m\to\infty} n(Q_{2^m\times 2^m}(R_k))/2^{2m} = \mu(R_k)$

$$\therefore \sum_{k=1}^{q}\mu(R_k) = \sum_{k=1}^{q}\lim_{2^m\to\infty} n(Q_{2^m\times 2^m}(R_k))/2^{2m} = \lim_{2^m\to\infty}\sum_{k=1}^{q} n(Q_{2^m\times 2^m}(R_k))/2^{2m}$$

$$\leq \lim_{2^m\to\infty} n(Q_{2^m\times 2^m}(\bigcup_{k=1}^{q} R_k))/2^{2m} \leq \lim_{2^m\to\infty} Q_{2^m\times 2^m}(B)/2^{2m}$$

$$\therefore \mu(B)-\varepsilon \leq \sum_{k=1}^{q}\mu(R_k) \leq \lim_{2^m\to\infty} Q_{2^m\times 2^m}(B)/2^{2m}$$

From the above, $\mu(B)-\varepsilon \leq \lim_{2^m\to\infty} n(Q_{2^m}(B))/2^m \leq \mu(B)$

$\varepsilon>0$ is arbitrary, so $\lim_{2^m\to\infty} Q_{2^m\times 2^m}(B)/2^{2m} = \mu(B)$

5. Solving Bertrand's (Chord) Paradox

5.1. The rigorous definition Bertrand's Problem

A chord is a line segment connecting two points on the circumference. Therefore, {a}×{b} ∈ S×S represents one chord, where a∈ S, b∈S, S is a circle. Of course, every chord represents twice in S×S, such as {a}×{b} and {b}×{a}.

Let S be a circle of radius $\frac{1}{2\pi}$, and let’s define a function f representing the length of a chord as follows :

$$f(x,y) = d(x,y)$$

(where $(x,y) \in$ S×S$\subset R^2 \times R^2$, $d(x,y)$ is Euclid distance,

if $x=(a,b)$, $y=(c,d)$ then $d(x,y)=\sqrt{(a-c)^2+(b-d)^2}$)

If 0 and 1 are defined as the same point in [0,1], then
[0,1]×[0,1] becomes the same topological space as S×S.
In S×S, the Lebesgue measure can also be defined in the same way as in [0,1]×[0,1].

The length of one side of an equilateral triangle inscribed in a circle of radius $\frac{1}{2\pi}$ is $\frac{\sqrt{3}}{2\pi}$.

Then, the set P that satisfies the assumption of Bertrand's Problem is defined as follows :

$$\text{P} = \{(x,y) \in \text{S×S} \mid d(x,y) > \frac{\sqrt{3}}{2\pi} \}$$

Now, Bertrand's Problem becomes a problem to find μ(P), where μ(P) is the Lebesgue measure of P.

5.2. The Applicability of J-Sequence m-collection cover counting of 2 dimension

Since $f(x,y) = d(x,y)$ is continuous,

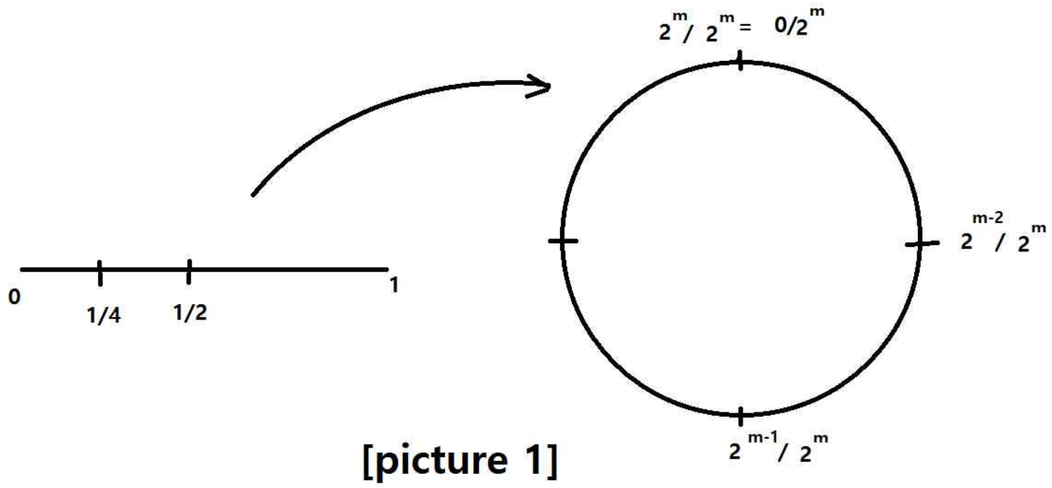
and set R={r| r > $\frac{\sqrt{3}}{2\pi}$ } is open,

then P = f^{-1}(R) is open.

Therefore, by Theorem 2, $\lim_{2^m \to \infty}$ n($Q_{2^m \times 2^m}$(P))/2^{2m}=μ(P), thus P is a set where J-Sequence m-collection cover counting of 2 dimension is applicable.

5.3. Probability calculation of set P

Calculation of $\lim_{2^m \to \infty}$ n($Q_{2^m \times 2^m}$(P))/2^{2m}, where n($Q_{2^m \times 2^m}$(P)) is J-sequence m-collection cover counting of 2 dimension about set P.



[picture 1]

Consider [0,1] as S like picture 1.

For given I_i and m, if there is a certain I_j such that for $\forall x \in I_i$ and $\forall y \in I_j$, $d(x,y)$ > $\frac{\sqrt{3}}{2\pi}$, then $I_i \times I_j \subset$ P, that is, $\mu(I_i \times I_j$-P)=0 and $I_i \times I_j \in Q_{2^m \times 2^m}$(P) .

Let p_i be the number of $I_i \times I_j$, where $I_i \times I_j \in$ $Q_{2^m \times 2^m}$(P),

then n($Q_{2^m \times 2^m}$(P)) = $\sum_{i=1}^{2^m} p_i$

Assume i = 1,

If [2^m/3]+3 ≤ j ≤ 2^m-[2^m/3]-1 ,

then for $\forall x \in I_1$ and $\forall y \in I_j$, $d(x,y)$ > $\frac{\sqrt{3}}{2\pi}$, thus $I_1 \times I_j$ $\in Q_{2^m \times 2^m}$(P).

($[x]$ is Gaussian function)

On the other hand,

if $j \leq [2^m/3]-1$ or $j \geq 2^m-[2^m/3]+3$,

then for $\forall x \in I_1$ and $\forall y \in I_j$, $d(x,y) < \frac{\sqrt{3}}{2\pi}$, that is, $(I_1 \times I_j) \cap \mathrm{P} = \varnothing$, thus $I_1 \times I_j \not\in Q_{2^m \times 2^m}(\mathrm{P})$.

Let $A_m = \{j \mid [2^m/3]+3 \leq j \leq 2^m-[2^m/3]-1\}$,

$B_m = \{j \mid j \leq [2^m/3]-1 \text{ or } j \geq 2^m-[2^m/3]+3\}$,

Since $p_1 \geq \mathrm{n}(A_m)$,

$$2^m-2[2^m/3]-3 \leq p_1$$

Since $p_1 \leq \mathrm{n}(B_m^C)$,

$$p_1 \leq 2^m-2[2^m/3]+3$$

$$\therefore\ 2^m-2[2^m/3]-3 \leq p_1 \leq 2^m-2[2^m/3]+3$$

For other i, rotating the circle makes I_1 and I_i coincide, so $p_i = p_1$

$$\therefore\ \mathrm{n}(Q_{2^m \times 2^m}(\mathrm{P})) = \sum_{i=1}^{2^m} p_i = p_1 \times 2^m$$

$$(2^m-2[2^m/3]-3) \times 2^m \leq \mathrm{n}(Q_{2^m \times 2^m}(\mathrm{P})) = p_1 \times 2^m \leq (2^m-2[2^m/3]+3) \times 2^m$$

$$\Leftrightarrow\ ((2^m-2[2^m/3]-3) \times 2^m)/2^{2m} \leq \mathrm{n}(Q_{2^m \times 2^m}(\mathrm{P}))/2^{2m} \leq ((2^m-2[2^m/3]+3) \times 2^m)/2^{2m}$$

$$\rightarrow\ \lim_{2^m \to \infty} \mathrm{n}(Q_{2^m \times 2^m}(\mathrm{P}))/2^{2m} = 1/3$$

$$\therefore\ \mu(\mathrm{P}) = \lim_{2^m \to \infty} \mathrm{n}(Q_{2^m \times 2^m}(\mathrm{P}))/2^{2m} = 1/3$$

5.4. The answer is 1/3.

6. Conclusion

When calculating probabilities, we have examined that several strange things may occur if not taking into account the structure of the infinite set in which σ -algebra is defined.

J-sequence shows a sampling method by considering the structure of the set in which σ-algebra is defined.

After the number of cases is calculated using the J-sequence m-collection cover, and if there exist its limit value, it becomes the probability of the event applying the number of cases.

For any sequence $\{a_k\}$, where $a_k \in [0,1]$, let $S_k = \{a_1, a_2, a_3, \ldots, a_k\}$.

Is there a sequence in which $\lim_{k\to\infty} \frac{1}{k}$ n(C∩S_k)/k = μ(C) for every measurable set C ?

This is related to the reason for counting by using the J-sequence m-collection cover, not the J-sequence itself. If such a sequence exists, any probability can be calculated by counting the number of cases using S_k as a sampling.

However, no such a sequence exists. Restricting C to be an open set also does not exist. For any sequence, we can produce an open set of sufficiently small measure which includes all of the sequence.

If we cover, for example, each element with an open interval of $\varepsilon/2^k$, then the unions of that intervals is an open set which contains all elements of the sequence and of which measure is less than ε.

A sequence that satisfies the above condition for every open interval can be found. That is an equidistributed sequence on [0,1]. ([8] Kuipers 2006:8)

It is that, for an equidistributed sequence $\{a_k\}$, $\lim_{k\to\infty} \frac{1}{k}$ n(A∩S_k)/k = b-a =μ(A) for every open interval A=(a,b)⊂[0,1].

Unfortunately, in the world of real numbers where infinity and infinity intersect, it can be seen that it is impossible to represent them equally with a sequence, that is, impossible to generate a sequence that, for all measurable sets, it is included in the set with a frequency proportional to the measure of the set.

Therefore, in a sample space composed of an uncountably infinite number of elements such as [0,1], after calculating the number of cases using sets generated by equally dividing the sample space into finite numbers, the probability of an event can be calculated with its limit value.

Reference

1. Bertrand, Joseph (1889), "Calcul des probabilités", Gauthier-Villars, p. 5-6.
2. Drory, Alon (2015), "Failure and Uses of Jaynes' Principle of Transformation Groups", Foundations of Physics, 45 (4): 439–460.
3. Shackel, N. (2007), "Bertrand's Paradox and the Principle of Indifference" (PDF), Philosophy of Science, 74 (2): 150–175,
4. Rudin, Walter (1987), Real & Complex Analysis. McGraw-Hill.
5. H. L. Royden(1989), Real Analysis, 3/e, Maxwell McMillan.
6. Clark, M(2012), Paradoxes from a to Z, 3/e. London: Routledge
7. Kolmogorov, A. N(1950), Foundations of the theory of probability. New York: Chelsea Pub. Co.
8. Kuipers, L.; Niederreiter, H.(2006). Uniform Distribution of Sequences. Dover Publications.