Introduction to probability and statistic: Assignment 1

Darix, SAMANI SIEWE

October 19, 2024

Exercise 1

I. Show that if an algebra (field) contains a finite number of sets, then it is a σ-algebra (σ-field).

A algebra **field** $\mathcal{F}$ of subsets of a sample space S is a collection of subsets such that:

(a) **Contains the whole space**: The sample space S is in F .
(b) **Closed under complementation**: If $A \in \mathcal{F}$, then $\bar{A} \in \mathcal{F}$.
(c) **Closed under finite unions**: If $A_i \in \mathcal{F}$, then $\bigcup_{i=1}^{n} F_i \in \mathcal{F}$. $\forall i \in 1 \ldots n$

A σ-algebra **σ-field** $\mathcal{F}$ of subsets of a sample space S is a collection of subsets such that:

(1) **Contains the whole space**: The sample space S is in F .
(2) **Closed under complementation**: If $A \in \mathcal{F}$, then $\bar{A} \in \mathcal{F}$.
(3) **Closed under infinite denombrable unions**: If $A_i \in \mathcal{F}$, then $\bigcup_{i=1}^{\infty} A_i \in \mathcal{F}$. $\forall i \in 1 \ldots n$

By definition the (a) and (b)(from the definition of algebra field) implies (1) and (2) (from the definition of σ-field). just missing to prove that the closed under infinite denombrable of unions is also in $\mathcal{F}$.

if the index set from (c)(from the definition of algebra field) is infinite: Since the collection of sets is finite, any countable union can only include a finite number of the sets from $\mathcal{F}$ F. Thus, if $\mathcal{A}_\rangle$ is chosen to be any set from the finite collection, the countable union will essentially still be one of the finite sets, since:

$\bigcup_{i=1}^{\infty} A_i = \mathcal{A}_|$ for some $j \in 1, \ldots, n$ ((if infinitely many sets are the same))

Since $\mathcal{F}$ is closed under finite unions and the infinite countable union of the sets can be expressed as a finite union of those sets, it follows that:

$\bigcup_{i=1}^{\infty} A_i \in \mathcal{F}$

if an algebra (field) contains a finite number of sets, then it is a σ-algebra (σ-field).

II. Let $\mathcal{F}$ be the collection of all subsets of $\omega = (0,1]$ of the type (a,b], where $a < b$ are rational numbers in (0,1] together with the empty set :

$\mathcal{F} = \{\emptyset, (a,b] \subseteq (0,1], a, b \in \mathbf{Q}, a < b\}$

1. Verify that $\mathcal{F}$ is an algebra (field).
   - $\Omega = (0,1] \subseteq (0,1] \implies \Omega \in \mathcal{F}$
   - let's denote by $A = (a,b]$ the complement of A is $\bar{A} = (0,a[\cup]a,b[$
     and we know that $(0,a) \in (0,1]$ and $(0,a) \in (0,1]$ and the union $(0,a[\cup]a,b[\in (0,1]$
   - let's denote by $V_i = (a_i, bi] \in \mathcal{F}$ let's prove that $\bigcup_{i=1}^{n} A_i \in \mathcal{F}$. $\forall i \in 1 \ldots n$
     we know that $V_i \in (0,1]$ and the union of all V_i is also include in $(0,1]$
     so the union of $\bigcup_{i=1}^{n} A_i \in \mathcal{F}$. $\forall i \in 1 \ldots n$
2. let $p_n = \dfrac{[10\pi]}{10^n}$ and $a_n = \dfrac{1}{p_n}$ where [x] means truncation to the nearest integer less than or equal to x.

(a) What is the limit of p_n when n goes to infinity ?

$$\begin{aligned}
10^n\pi \le [10^n\pi] &< 10^n\pi + 1 \\
\Longrightarrow \frac{10^n\pi}{10^n} \le \frac{[10^n\pi]}{10^n} &< \frac{10^n\pi+1}{10^n} \\
\Longrightarrow \pi \le \frac{[10^n\pi]}{10^n} &< \pi + \frac{1}{10^n}
\end{aligned}$$

$\lim\limits_{n\to\infty}\left(\pi + \frac{1}{10^n}\right) = \pi$ and by squeeze theorem:

$\lim\limits_{n\to\infty}\left(\frac{10^n\pi+1}{10^n}\right) = \pi$

(b) Consider the sequence of intervals of the form $(a_n, 1]$ $n = 1, \ldots, n$ from $\mathcal{F}$. does $\bigcup\limits_{i=1}^{\infty}(a_n, 1] \in \mathcal{F}$ belong $\mathcal{F}$.

to answer to this question let's following the step: find monotonic of p_n after deduce monotonic of a_n, and finally find the interval of $(a_i, 1]$ that include all other interval by using the monotonic of a_n find the monotonic of p_n:

$\frac{p_{n+1}}{p_n} = \frac{\frac{[10^{n+1}\pi]}{10^{n+1}}}{\frac{[10^n\pi]}{10^n}} = \frac{[10^{n+1}\pi]}{10^{n+1}} * \frac{10^n}{[10^n\pi]} = \frac{[10^{n+1}\pi]}{10[10^n\pi]}$

$$10^{n+1}\pi \le [10^{n+1}\pi] < 10^{n+1}\pi + 1 \qquad (1)$$

and

$$\begin{aligned}
10^n\pi \le [10^n\pi] &< 10^n\pi + 1 \\
\Longrightarrow \frac{1}{10^n\pi+1} < \frac{1}{[10^n\pi]} &\le \frac{1}{10^n\pi} \\
\Longrightarrow \frac{1}{10^n\pi+1} < \frac{1}{[10^n\pi]} &\le \frac{1}{10^n\pi} \\
\Longrightarrow \frac{1}{10(10^n\pi+1)} < \frac{1}{10[10^n\pi]} &\le \frac{1}{10*10^n\pi} \qquad (2)
\end{aligned}$$

(1) * (2) we have,

$$\begin{aligned}
\frac{10^{n+1}\pi}{10(10^n\pi+1)} < \frac{[10^{n+1}\pi]}{10[10^n\pi]} &< \frac{10^{n+1}\pi+1}{10*10^n\pi} \\
\frac{10^{n+1}\pi}{10(10^n\pi+1)} < \frac{[10^{n+1}\pi]}{10[10^n\pi]} &< \frac{10^{n+1}\pi+1}{10^{n+1}\pi}
\end{aligned}$$

$\frac{p_{n+1}}{p_n} = \frac{[10^{n+1}\pi]}{10[10^n\pi]} > \frac{10^{n+1}\pi}{10*10^n\pi} > 1$

so p_n is a decreasing sequence and that means that a_n is a decreasing sequence,

so $\bigcup\limits_{i=1}^{n}(a_n, 1] \subseteq (a_\infty, 1] = (\frac{1}{\pi}, 1] \subseteq (0, 1]$

based on prove bellow we can assume that $\bigcup\limits_{i=1}^{\infty}(a_n, 1] \in \mathcal{F}$ is belong $\mathcal{F}$.

(c) Write a conclusion.

based on the preview question, we have show that $\mathcal{F}$ is a $\sigma-$field

Exercise 2

Consider an experiment with the following sample space

$\Omega = \{abc, bac, cab, bca, acb, cba, aaa, bbb, ccc\}$

where each outcome is equally likely. Define events A_k = '$a$in k-th position', for k = 1,2,3.

1. Describe each event $A_k, k = 1, 2, 3$.

   $A_1 = \{abc, acb, aaa, \}$
   $A_2 = \{bac, cab, aaa, \}$
   $A_3 = \{bca, cba, aaa, \}$

2. Are these events mutually independent ?

   $A_1 \cap A_2 \cap A_3 = \{aaa\}$

   $P(A_1 \cap A_2 \cap A_3) = \frac{Card(A_1 \cap A_2 \cap A_3)}{Card(\Omega)} = \frac{1}{9}$ and

   $P(A_1) * P(A_2) * P(A_3)) = \frac{3}{9} * \frac{3}{9} * \frac{3}{9} = \frac{1}{3} * \frac{1}{3} * \frac{1}{3} = \frac{1}{27}$

   so, $P(A_1) * P(A_2) * P(A_3)) \neq P(A_1 \cap A_2 \cap A_3)$

   so these events are not mutually independent

3. Are they independent events ?

   to answer of this question we need to check if these events are independent two by two

   - check for A_1 and A_2 :
     $P(A_1) * P(A_2) = \frac{3}{9} * \frac{3}{9} = \frac{1}{3} * \frac{1}{3} = \frac{1}{9}$ and $P(A_1 \cup A_2) = \frac{Card()A_1 \cup A_2}{Card(\Omega)} = \frac{1}{9}$
     so $P(A_1) * P(A_2) = P(A_1 \cup A_2) = \frac{1}{9}$
     the event A_1 and A_2 are independent
   - check for A_1 and A_3 :
     $P(A_1) * P(A_3) = \frac{3}{9} * \frac{3}{9} = \frac{1}{3} * \frac{1}{3} = \frac{1}{9}$ and $P(A_1 \cup A_3) = \frac{Card()A_1 \cup A_3}{Card(\Omega)} = \frac{1}{9}$
     so $P(A_1) * P(A_3) = P(A_1 \cup A_3) = \frac{1}{9}$
     the event A_1 and A_3 are independent
   - check for A_2 and A_3 :
     $P(A_2) * P(A_3) = \frac{3}{9} * \frac{3}{9} = \frac{1}{3} * \frac{1}{3} = \frac{1}{9}$ and $P(A_2 \cup A_3) = \frac{Card()A_1 \cup A_3}{Card(\Omega)} = \frac{1}{9}$
     so $P(A_2) * P(A_3) = P(A_1 \cup A_2) = \frac{1}{9}$
     the event A_2 and A_3 are independent

   Based on the result bellow we can assume that these events are independent

Exercise 3

Imagine a population of N +1 urns. Urn number k contains k red balls and $N - k$ white balls ($k = 0, 1, \ldots, N$). An urn is chosen at random and n random drawings are made from it, the ball drawn being replaced each time.
Define
Event A : all balls turn out to be red.
Event B : the (n+1)st draw yields a red ball.

1. Find $P(A|\text{Urn k is chosen})$, (k = 0,1,..., N).

   To answer this question, let's calculate the case by case and then generalize the expression depending on K. and suppose that for this exercise we have an equiprobability of choosing an Urn and we know that by the definition of equiprobabiliy $p(U_k) = \frac{1}{N+1}$

   if U_k is given the probably of A in the first time is $\frac{card(A)}{Card(\Omega)} = \frac{k}{N}$ and we know that the probability of frist time choosing is independent with the second time choosing and another time (n time) so $P(A|U_k) = \left(\frac{k}{N}\right)^n$

   - case when k = 0:
     $P(A|U_0) = 0$
   - case when k = 1:
     $P(A|U_1) = \left(\frac{1}{N}\right)^n = \left(\frac{1}{N}\right)^n$
   - case when k = 2:
     $P(A|U_2) = \left(\frac{2}{N}\right)^n = \left(\frac{1}{N}\right)^n$

   so for the generalization, $P(A|U_k) = \left(\frac{k}{N}\right)^n$

2. Find P(A), $P(A \cap B)$ and $P(B|A)$.

   using the probability total we can find the expression of P(A because) $\cup U_k$ $k = 0, 1, \ldots, N+1$ is a finite or countably infinite set of mutually exclusive and collectively exhaustive events

$$\begin{aligned} P(A) &= \sum_{k=0}^{N} P(A|U_i) * P(U_k) \\ &= \sum_{k=0}^{N} P(A|U_k) * \frac{1}{N+1} \\ &= \frac{1}{N+1} \sum_{k=0}^{N} \left(\frac{k}{N}\right)^n \end{aligned}$$

   so $P(A) = \frac{1}{N+1} \sum_{k=0}^{N} \left(\frac{k}{N}\right)^n$

   $A \cap B$: all (n+1) first random drawings is red. it is just A when n is replace by (n+1)

   $P(A \cap B) = \frac{1}{N+1} \sum_{k=0}^{N} \left(\frac{k}{N}\right)^{n+1}$

   by using the Bayes's theorem: we can able to express $P(B|A)$ easily :

   $P(B|A) = \dfrac{P(A \cap B)}{P(A)} = \dfrac{\frac{1}{N+1} \sum_{k=0}^{N} \left(\frac{k}{N}\right)^{n+1}}{\frac{1}{N+1} \sum_{k=0}^{N} \left(\frac{k}{N}\right)^{n}} = \dfrac{\sum_{k=0}^{N} \left(\frac{k}{N}\right)^{n+1}}{\sum_{k=0}^{N} \left(\frac{k}{N}\right)^{n}} = \dfrac{\sum_{k=1}^{N} \left(\frac{k}{N}\right)^{n+1}}{\sum_{k=1}^{N} \left(\frac{k}{N}\right)^{n}}$

3. Find an approximation to $P(B|A)$, using the fact that if N is large,

$$\begin{aligned} P(B|A) &= \frac{\sum_{k=1}^{N} \left(\frac{k}{N}\right)^{n+1}}{\sum_{k=1}^{N} \left(\frac{k}{N}\right)^{n}} \\ &= \frac{N \int_0^1 x^{n+1} dx}{N \frac{1}{n+1}} \\ &= \frac{N \frac{1}{n+2}}{N \frac{1}{n+1}} \\ &= \frac{\frac{1}{n+2}}{\frac{1}{n+1}} \\ &= \frac{n+1}{n+2} \end{aligned}$$

   so when N is very large, $P(B|A) = \frac{n+1}{n+2} = 1 - \frac{1}{n+2}$

   Interpretation of this Result

   This means that if the first n random drawings are red, the probability that the $(n+1)^{th}$ random drawing will also be red approaches 1. In practical terms, as we continue to draw from a large population, the occurrence of consecutive red draws significantly increases the likelihood that the next draw will be red as well. Thus, after observing many red outcomes, we become increasingly confident that the trend will continue.

Exercise 3

Let X be a **geometric** random variable with parameter p.

$P(X = x) = (1-p)^x p \; 0 < p < 1, x \in \mathbb{N}$

1. Calculate the moment generating function and deduce the expectation of X.

$$
\begin{aligned}
M_x(t) = E(e^{Xt}) &= \sum_{x\in\mathbb{N}} e^{xt} P(X = x) \\
&= \sum_{x\in\mathbb{N}} e^{xt}(1-p)^x p \\
&= p\sum_{x\in\mathbb{N}} (e^t(1-p))^x \\
&= p \lim_{N\to\infty} \frac{1-(e^t(1-p))^{N+1}}{1-e^t(1-p)} \\
&= \frac{p}{1-e^t(1-p)} \qquad\qquad t < -\ln(1-p)
\end{aligned}
$$

$M_x(t) = \dfrac{p}{1-e^t(1-p)} \qquad\qquad t < -\ln(1-p)$

now lets' deduce the expectation

$$
M_X^{'}(t) = \frac{(1-p)pe^t}{(1-e^t(1-p))^2}
$$

$$
\begin{aligned}
E(X) = M_X^{'}(0) &= \frac{(1-p)p}{p^2} \\
&= \frac{(1-p)}{p}
\end{aligned}
$$

so, $E(X) = \dfrac{(1-p)}{p}$

2. Let Y be another geometric r.v. with parameter p. Suppose X and Y are independent. Calculate the moment generating function of X +Y.

$$
\begin{aligned}
M_{X+Y}(t) = M_X(t)M_Y(t) &= \frac{p}{1-e^t(1-p)} * \frac{p}{1-e^t(1-p)} \\
&= \frac{p^2}{(1-e^t(1-p))^2}
\end{aligned}
$$

so $M_{X+Y}(t) = \dfrac{p^2}{(1-e^t(1-p))^2}$

3. Let Z be a **negative-binomiale** r.v. with parameters (n, θ) :

   $P(Z = z) = \dbinom{n+z-1}{z}\theta^n(1-\theta)^z \; z \in \mathbb{N}, 0 < \theta < 1$

   (a) Show that this is a probability.

   $\forall z \in \mathbb{N}, 0 < \theta < 1, P(Z = z) > 0$
   and

$$
\begin{aligned}
\sum_{z\in\mathbb{N}} P(Z=z) &= \sum_{z\in\mathbb{N}} \binom{n+z-1}{z}\theta^n(1-\theta)^z \\
&= \theta^n \sum_{z=0N}^{\infty} \binom{n+z-1}{z}(1-\theta)^z \\
&= \frac{\theta^n}{(1-(1-\theta))^n} \\
&= \frac{\theta^n}{\theta^n} \\
&= 1
\end{aligned}
$$

so based on the preview result this distribution is a probability

(b) Express the moment generating function of Z.

$$
\begin{aligned}
M_z(t) = E(e^{tx}) &= \sum_{z\in\mathbb{N}} e^{zt}\binom{n+z-1}{z}\theta^n(1-\theta)^z \\
&= \theta^n \sum_{z\in\mathbb{N}} e^{zt}\binom{n+z-1}{z}(1-\theta)^z \\
&= \theta^n \sum_{z\in\mathbb{N}} \binom{n+z-1}{z}(e^t(1-\theta))^z \\
&= \frac{\theta^n}{(1-e^t(1-\theta))^n}
\end{aligned}
$$

so, $M_z(t) = \dfrac{\theta^n}{(1-e^t(1-\theta))^n}$

4. Deduce the distribution of X +Y.

   based on the result of 2 and 3 (b) the variable X+Y follow the distribution **negative-binomial** with parameter (2, p)

5. Generalize this result to the summation of n independent geometric r.v. with parameter p.

   let's denote by $S_n = \sum_{i=1}^{n} X_i$ where X_i the summation of n independent geometric distribution variable

   $M_{S_n} = \sum_{i=1}^{n} M_{X_i} = \sum_{i=1}^{n} \frac{p}{1-e^t(1-p)} = \frac{p^n}{(1-e^t(1-p))^n}$

   based on this result the summation of n independent geometric distribution variable follows the negative-binomial with parameters (n, p)