Combinatorics: classical questions around the cardinal of random sets

The goal of this short problem is to derive classical results concerning subsets chosen uniformly at random from a given finite set. We begin with a combinatorial approach and then show how introducing random variables can simplify the analysis and reduce the need for explicit computations.

Preliminaries

1. Prove the binomial theorem:

$$\forall x, y \in \mathbb{R}, \forall n \in \mathbb{N}, (x+y)^n = \sum_{k=0}^{n} C_n^k x^k y^{n-k}.$$

2. Deduce that $\forall x, y \in \mathbb{R}, \forall n \in \mathbb{N}$,

$$\sum_{k=0}^{n} k C_n^k x^k y^{n-k} = nx(x+y)^{n-1} \quad \text{and} \quad \sum_{k=0}^{n} k^2 C_n^k x^k y^{n-k} = nx(nx+y)(x+y)^{n-2}.$$

For any set E, we call power set of E the set $\mathcal{P}(E)$ of the subsets of the set E.

In all what follows, we consider a natural number $n \in \mathbb{N}$.

Combinatorics in the power set – the basics

3. Show that for any $k \in \{0, \ldots, n\}$, the number of elements of $\mathcal{P}(\{1, \ldots, n\})$ of cardinal k is C_n^k.

4. Deduce that the cardinal of $\mathcal{P}(\{1, \ldots, n\})$ is 2^n.

We consider a random variable $\tilde{A}$ uniformly distributed in $\mathcal{P}(\{1, \ldots, n\})$, i.e. for all $A \subset \{1, \ldots, n\}$, $\mathbb{P}(\tilde{A} = A) = \frac{1}{2^n}$.

5. Compute $\mathbb{E}[\text{Card}(\tilde{A})]$ and $\mathbb{V}[\text{Card}(\tilde{A})]$.

Combinatorics in the power set – intersection and union

We now consider two independent random variables $\tilde{A}$ and $\tilde{B}$ uniformly distributed in $\mathcal{P}(\{1, \ldots, n\})$.

6. Let $m \in \mathbb{N}$. Show that if $\text{Card}E = m$, then the number of pairs $(F, G) \in \mathcal{P}(E)^2$ such that $F \cap G = \emptyset$ is $\sum_{j=0}^{m} C_m^j 2^{m-j} = 3^m$.

7. Let $k \in \{0, \ldots, n\}$. Show that the number of pairs $(A, B) \in \mathcal{P}(\{1, \ldots, n\})^2$ such that $A \cap B$ has exactly k elements is $C_n^k 3^{n-k}$. Deduce the value of $\mathbb{P}(\text{Card}(\tilde{A} \cap \tilde{B}) = k)$.

8. Deduce that $\mathbb{E}[\text{Card}(\tilde{A} \cap \tilde{B})] = \frac{n}{4}$ and $\mathbb{V}[\text{Card}(\tilde{A} \cap \tilde{B})] = \frac{3n}{16}$.

9. Deduce that $\mathbb{E}[\mathrm{Card}(\tilde{A} \cup \tilde{B})] = \frac{3n}{4}$ and $\mathbb{V}[\mathrm{Card}(\tilde{A} \cup \tilde{B})] = \frac{3n}{16}$.
Hint: consider the random variables $\{1, \ldots, n\} \setminus \tilde{A}$ *and* $\{1, \ldots, n\} \setminus \tilde{B}$.

Combinatorics in the power set – inclusion

We continue to consider two independent random variables $\tilde{A}$ and $\tilde{B}$ uniformly distributed in $\mathcal{P}(\{1, \ldots, n\})$.

10. Let $k \in \{0, \ldots, n\}$. Show that the number of pairs $(A, B) \in \mathcal{P}(\{1, \ldots, n\})^2$ such that $\mathrm{Card}(A) = k$ and $A \subset B$ is $C_n^k 2^{n-k}$.

11. Deduce that the number of pairs $(A, B) \in \mathcal{P}(\{1, \ldots, n\})^2$ such that $A \subset B$ is 3^n.

12. Deduce the value of $\mathbb{P}(\tilde{A} \subset \tilde{B})$.

13. What is the probability of the event "Among $\tilde{A}$ and $\tilde{B}$, one is included into the other"?

Boolean variables

We continue to consider two independent random variables $\tilde{A}$ and $\tilde{B}$ uniformly distributed in $\mathcal{P}(\{1, \ldots, n\})$.

We introduce $2n$ independent Bernoulli random variables $X_1, \ldots, X_n, Y_1, \ldots, Y_n$ with parameter $\frac{1}{2}$.

We define two random variables $\tilde{A}'$ and $\tilde{B}'$ with values in $\mathcal{P}(\{1, \ldots, n\})$ by:

$$\forall i \in \{1, \ldots, n\}, \quad i \in \tilde{A}' \iff X_i = 1 \quad \text{and} \quad i \in \tilde{B}' \iff Y_i = 1.$$

14. Prove that $\tilde{A}'$ and $\tilde{B}'$ are well defined.

15. Prove that $(\tilde{A}', \tilde{B}')$ has the same distribution as $(\tilde{A}, \tilde{B})$.

16. Deduce that $\mathrm{Card}(\tilde{A} \cap \tilde{B})$ and $\sum_{i=1}^n X_i Y_i$ have the same distribution.

17. Conclude that $\mathrm{Card}(\tilde{A} \cap \tilde{B})$ follows a binomial distribution $\mathcal{B}\left(n, \frac{1}{4}\right)$.

18. Using a similar reasoning, prove that $\mathrm{Card}(\tilde{A} \cup \tilde{B})$ follows a binomial distribution $\mathcal{B}\left(n, \frac{3}{4}\right)$.

19. Deduce the first two moments of $\mathrm{Card}(\tilde{A} \cap \tilde{B})$ and $\mathrm{Card}(\tilde{A} \cup \tilde{B})$. Comment.

20. Prove that $\mathbb{P}(\tilde{A} \subset \tilde{B}) = \mathbb{P}\left(\bigcap_{i=1}^n \{X_i \leq Y_i\}\right)$.

21. Deduce that $\mathbb{P}(\tilde{A} \subset \tilde{B}) = \left(\frac{3}{4}\right)^n$. Comment.