

# My probability and statistics exercises

Evgeny Markin

2023

# Contents

<b>1</b>	<b>Introduction to Probability</b>	<b>2</b>
1.1	The History of Probability . . . . .	2
1.2	Interpretations of Probability . . . . .	2
1.3	Experiments and Events . . . . .	2
1.4	Set Theory . . . . .	2
1.5	The Definition of Probability . . . . .	2
1.6	Finite Sample Spaces . . . . .	5
1.7	Counting Methods . . . . .	6
1.8	Combinatorial Methods . . . . .	7
1.9	Multinomial Coefficients . . . . .	12
1.10	The Probability of a Union of Events . . . . .	12

# Chapter 1

## Introduction to Probability

### 1.1 The History of Probability

### 1.2 Interpretations of Probability

### 1.3 Experiments and Events

### 1.4 Set Theory

*Exercises in this section (or exercises similar to them) are handled in the set theory course*

### 1.5 The Definition of Probability

1	$2/5$
2	0.7
3a	$1/2$
3b	$1/6$
3c	$3/8$
4	0.6
5	0.4
6	0.5
8	30
11a	$1 - \pi/4$
11b	0.75
11c	$2/3$
11d	0
14a	0.38, 0.16
14b	0.04

A little notation, related to 6:

$$Pr(A) = 0.5$$

$$Pr(B) = 0.2$$

$$Pr(A \cap B) = 0.1$$

$$Pr(A \cup B) = 0.6$$

$$Pr((A \cup B) \cap (A \cap B)^c) = P(A \cup B) - P((A \cup B) \cap (A \cap B)) = P(A \cup B) - P(A \cap B) = 0.5$$

### 1.5.7

If  $Pr(A) = 0.4$  and  $Pr(B) = 0.7$ , then we follow that the maximum  $Pr(A \cap B)$  is attained if  $A \subset B$ , in which case  $Pr(A \cap B) = Pr(A) = 0.4$ . The minimum is obtained if  $A \cup B = S$ , in which case  $Pr(A \cap B) = 0.1$

### 1.5.9

The event that exactly one of the events occurs can be expressed as

$$(A \cap B^c) \cup (A^c \cap B)$$

which comes from either the definition of xor, common sense or something else, depending on your preferences. Thus we follow that

$$\begin{aligned} Pr((A \cap B^c) \cup (A^c \cap B)) &= Pr(A \cap B^c) + Pr(A^c \cap B) - Pr((A \cap B^c) \cap (A^c \cap B)) = \\ &= Pr(A \cap B^c) + Pr(A^c \cap B) - Pr((A \cap A^c) \cap (B^c \cap B)) = \\ &= Pr(A \cap B^c) + Pr(A^c \cap B) = Pr(A) - Pr(A \cap B) + Pr(B) - Pr(B \cap A) = \\ &= Pr(A) - Pr(A \cap B) + Pr(B) - Pr(A \cap B) = Pr(A) + Pr(B) - 2Pr(A \cap B) \end{aligned}$$

as desired (rules used in this derivation: association of unions,  $A \cap A^c = \emptyset$  and other trivial stuff)

### 1.5.10

$$Pr(A \cap B^c) = Pr(A) - Pr(A \cap B)$$

$$Pr(A \cap B^c) + Pr(A \cap B) = Pr(A)$$

as desired.

**1.5.12**

Suppose that  $n > m \in N$ . Then we follow that by definition

$$B_m \subseteq A_m$$

and

$$B_n \subseteq A_m^c$$

thus we follow that

$$B_m \cap B_n \subseteq A_m \cap A_m^c = \emptyset$$

thus

$$B_m \cap B_n = \emptyset$$

therefore we conclude that  $B_1, B_2, \dots$  are disjoint sets. Thus we follow that

$$Pr(\bigcup_{i=1}^n B_i) = \sum_{i=1}^n Pr(B_i)$$

For  $n = 2$  we've got that

$$B_1 \cup B_2 = A_1 \cup (A_1^c \cap A_2) = (A_1 \cup A_1^c) \cap (A_1 \cup A_2) = A_1 \cup A_2$$

and by induction we can follow that

$$\bigcup_{i=1}^n B_i = \bigcup_{i=1}^n A_i$$

thus

$$Pr(\bigcup_{i=1}^n B_i) = \sum_{i=1}^n Pr(B_i)$$

implies that

$$Pr(\bigcup_{i=1}^n A_i) = \sum_{i=1}^n Pr(B_i)$$

for  $n \in N$ . Given that  $n$  is arbitrary, we can follow that

$$Pr(\bigcup_{i=1}^{\infty} A_i) = \sum_{i=1}^{\infty} Pr(B_i)$$

as desired.

**1.5.13**

First equation follow from induction on the result that

$$Pr(A \cup B) \leq Pr(A) + Pr(B)$$

the second equation follows from the first equation, DeMorgan laws and induction on the form

$$Pr(A \cap B) = Pr((A^c \cup B^c)^c) = 1 - Pr(A^c \cup B^c) \geq 1 - (Pr(A^c) + Pr(B^c))$$

**1.5.14**

$$Pr(A) = 0.34$$

$$Pr(B) = 0.12$$

$$Pr(O) = 0.5$$

$$Pr(AB) = 1 - 0.34 - 0.12 - 0.5 = 0.04$$

$$Pr(a - A) = 0.34 + 0.04 = 0.38$$

$$Pr(a - B) = 0.12 + 0.04 = 0.16$$

**1.6 Finite Sample Spaces**

1	1/2
2	1/2
3	2/3
4	1/7
5	4/7
6	1/4
8b	1/4

**1.6.7**

The possible genotypes are  $Aa$  and  $aa$  with probabilities  $1/2$  and  $1/2$  respectively

**1.6.8a**

The sample space of the experiment is  $\{heads, tails\} \times \{1, 2, 3, 4, 5, 6\}$ ,

## 1.7 Counting Methods

1	14
2	9000
3	120
4	24
5	5/18
6	5/324
7	0.014731
8	360 / 2401
9	1 / 20
10a	r/100
10b	r/100
10c	r/100

### 1.7.11

$$s(n) = \frac{1}{2} \log(2\pi) + (n + \frac{1}{2}) \log n - n \approx \log n!$$

$$\log n! - \log(n-m)! = \log \frac{n!}{(n-m)!}$$

$$\begin{aligned} s(n) - s(n-m) &= \frac{1}{2} \log(2\pi) + (n + \frac{1}{2}) \log n - n - (\frac{1}{2} \log(2\pi) + ((n-m) + \frac{1}{2}) \log(n-m) - (n-m)) = \\ &= (n + \frac{1}{2}) \log n - n - ((n-m) + \frac{1}{2}) \log(n-m) + (n-m) = \\ &= (n + \frac{1}{2}) \log n - ((n-m) + \frac{1}{2}) \log(n-m) - m \approx \log \frac{n!}{(n-m)!} \end{aligned}$$

$$P(n, m) = \frac{n!}{(n-m)!} = \exp(s(n) - s(n-m))$$

## 1.8 Combinatorial Methods

1	184756
2	latter
3	equal
4	1 / 10626
5	-
6	2/n
7	(n - k - 1)/C(n, k)
8	(n - k)/C(n, k)
9	(n + 1)/C(2n, n)
10	15/92 $\approx$ 0.16304
11	1/75 $\approx$ 0.01333
12	69/119 $\approx$ 0.57983
13	173/1518 $\approx$ 0.114
14	-
15	-
16a	48/175 $\approx$ 0.27429
16b	$2^{50}/C(100, 50) \approx 0$
17	$4C(13, 4)/C(52, 4) = 44/4165 \approx 0.0105$
18	$C(20, 2)^5/C(100, 10) \approx 0.0143$
19	-
20	-
21	$C(365 + 7 - 1, 7)$
22	-

### 1.8.5

*Prove that*

$$\frac{\prod_{4155 \leq i \leq 4251} i}{\prod_{2 \leq i \leq 97} i}$$

*is an integer*

$$\begin{aligned} & \frac{\prod_{4155 \leq i \leq 4251} i}{\prod_{2 \leq i \leq 97} i} = \frac{\prod_{4155 \leq i \leq 4251} i}{\prod_{1 \leq i \leq 97} i} = \\ &= \frac{\prod_{4155 \leq i \leq 4251} i}{97!} = \frac{4251!}{4154!97!} = \frac{4251!}{4154!(4251 - 4174)!} = C(4251, 4154) \end{aligned}$$

and binomial coefficients are integers (pretty sure that we can follow that by induction in some more advanced course).



**1.8.10**

There are total of  $C(24, 10)$  possible subsets of length 10 in the space of 24. We follow that there are  $C(22, 8)$  ways to pick 8 normal bulbs, which is what required to pick 2 defective bulbs. Therefore the probability is

$$\frac{C(22, 8)}{C(24, 10)} = 15/92 \approx 0.16304...$$

**1.8.12**

Using the same logic as in 1.8.10, there is a possibility  $\frac{C(33, 8)}{C(35, 10)}$  that same two guys will be in the first team, and probability of  $\frac{C(33, 23)}{C(35, 10)}$  that they'll be in the other team. Thus the total probability is the sum of two.

**1.8.14**

*Prove that for all positive integers  $n, k$  such that  $n \geq k$*

$$C(n, k) + C(n, k - 1) = C(n + 1, k)$$

$$\begin{aligned} C(n, k) + C(n, k - 1) &= \frac{n!}{(n - k)!k!} + \frac{n!}{(n - k + 1)!(k - 1)!} = \\ &= \frac{n!}{k(n - k)!(k - 1)!} + \frac{n!}{(n - k + 1)(n - k)!(k - 1)!} = \\ &= \frac{(n - k + 1)n!}{k(n - k + 1)(n - k)!(k - 1)!} + \frac{kn!}{k(n - k + 1)(n - k)!(k - 1)!} = \\ &= \frac{(n - k + 1)n! + kn!}{k(n - k + 1)(n - k)!(k - 1)!} = \frac{n!((n - k + 1) + k)}{k(n - k + 1)(n - k)!(k - 1)!} = \\ &= \frac{n!(n + 1)}{k(n - k + 1)(n - k)!(k - 1)!} = \frac{(n + 1)!}{((n + 1) - k)!k!} = C(n + 1, k) \end{aligned}$$

as desired.

**1.8.15**

*(a) Prove that*

$$\sum_{i=0}^n C(n, i) = 2^n$$

We can follow that from the fact that there are  $2^n$  subsets of any given finite set, which means that the number of subsets of different lengths sums up to  $2^n$ .

Another way to do this is to use binomial theorem:

$$(x + y)^n = \sum_{i=0}^n C(n, i)x^i y^{n-i}$$

thus if we substitute  $x$  and  $y$  for 1, we get

$$(1 + 1)^n = \sum_{i=0}^n C(n, i)1^i 1^{n-i}$$

$$2^n = \sum_{i=0}^n C(n, i)$$

(b) *Prove that*

$$\sum_{i=0}^n (-1)^i C(n, i) = 0$$

I'm sure that there is a neat explanation for this one as well, but using the binomial theorem once again, but now substituting 1 for  $x$  and  $-1$  for  $y$  we get

$$(1 - 1)^n = \sum_{i=0}^n C(n, i)1^i (-1)^{n-i}$$

$$\sum_{i=0}^n C(n, i)1^i (-1)^{n-i} = 0$$

we can follow through the even-odd argument that  $1^i (-1)^{n-i} = (-1)^i$ , but I'll skip it.

### 1.8.19

(rewording) *Prove the formula for unordered sampling with replacement.*

This thing is ought to be covered rigorously in a course for discrete maths, combinatorics or something of sorts. Currently there is a better proof at Belcastro's "Discrete mathematics with ducks".

### 1.8.20

*Prove the binomial theorem 1.8.2*

1.8.2 states that

$$(x + y)^n = \sum_{i=0}^n C(n, i)x^i y^{n-i}$$

Let

$$I = \{n \in \omega : (x + y)^n = \sum_{i=0}^n C(n, i)x^i y^{n-i}\}$$

We follow that

$$(x + y)^0 = C(0, 0)x^0 y^0 = 1$$

Thus  $0 \in I$ . (we can start with a base case of 1 as well for a more clear example, but I like this one more, and it suffices as well).

Now suppose that  $n \in I$ . We follow that

$$(x + y)^n = \sum_{i=0}^n C(n, i)x^i y^{n-i}$$

thus we follow that

$$(x + y)(x + y)^n = (x + y) \left[ \sum_{i=0}^n C(n, i)x^i y^{n-i} \right]$$

Left-hand side is reduced to

$$(x + y)(x + y)^n = (x + y)^{n+1}$$

Right-hand side is obviously a bit trickier, but we can follow

$$\begin{aligned} (x + y) \sum_{i=0}^n C(n, i)x^i y^{n-i} &= \\ &= x \sum_{i=0}^n C(n, i)x^i y^{n-i} + y \sum_{i=0}^n C(n, i)x^i y^{n-i} = \\ &= \sum_{i=0}^n C(n, i)x^{i+1} y^{n-i} + \sum_{i=0}^n C(n, i)x^i y^{n+1-i} = \\ &= \sum_{i=0}^n C(n, i)x^i y^{n+1-i} + \sum_{i=0}^n C(n, i)x^{i+1} y^{n-i} = \\ &= C(n, n)x^{n+1} y^0 + \sum_{i=0}^n C(n, i)x^i y^{n+1-i} + \sum_{i=0}^{n-1} C(n, i)x^{i+1} y^{n-i} = \\ &= x^{n+1} + \sum_{i=0}^n C(n, i)x^i y^{n+1-i} + \sum_{i=0}^{n-1} C(n, i)x^{i+1} y^{n-i} = \end{aligned}$$

$$\begin{aligned}
&= x^{n+1} + \sum_{i=0}^n C(n, i) x^i y^{n+1-i} + x \sum_{i=0}^{n-1} C(n, i) x^i y^{n-i} = \\
&= x^{n+1} + \sum_{i=0}^n C(n, i) x^i y^{n+1-i} + x \sum_{i=1}^n C(n, i-1) x^{i-1} y^{n-(i-1)} = \\
&= x^{n+1} + C(n, 0) x^0 y^{n+1} + \sum_{i=1}^n C(n, i) x^i y^{n+1-i} + \sum_{i=1}^n C(n, i-1) x^i y^{n+1-i} = \\
&= x^{n+1} + y^{n+1} + \sum_{i=1}^n C(n, i) x^i y^{n+1-i} + \sum_{i=1}^n C(n, i-1) x^i y^{n+1-i} = \\
&= x^{n+1} + y^{n+1} + \sum_{i=1}^n (C(n, i) + C(n, i-1)) x^i y^{n+1-i} = \\
&= x^{n+1} + y^{n+1} + \sum_{i=1}^n C(n+1, i) x^i y^{n+1-i} = x^{n+1} + C(n+1, 0) x^0 y^{n+1-0} + \sum_{i=1}^n C(n+1, i) x^i y^{n+1-i} = \\
&= x^{n+1} + \sum_{i=0}^n C(n+1, i) x^i y^{n+1-i} = x^{n+1} y^0 + \sum_{i=0}^n C(n+1, i) x^i y^{n+1-i} = \\
&= C(n+1, n+1) x^{n+1} y^{n+1-(n+1)} + \sum_{i=0}^n C(n+1, i) x^i y^{n+1-i} = \sum_{i=0}^{n+1} C(n+1, i) x^i y^{n+1-i}
\end{aligned}$$

Thus we follow

$$(x + y)^{n+1} = \sum_{i=0}^{n+1} C(n+1, i) x^i y^{n+1-i}$$

or

$$(x + y)^{n^+} = \sum_{i=0}^{n^+} C(n^+, i) x^i y^{n^+-i}$$

which means that  $n \in I \Rightarrow n^+ \in I$ , from which we conclude that  $I = \omega$ , and thus

$$(x + y)^n = \sum_{i=0}^n C(n, i) x^i y^{n-i}$$

for all  $n \in \omega$ , as desired.

### 1.8.22

Skip

## 1.9 Multinomial Coefficients

1	$(21!)/(7! * 7! * 7!)$
2	$50!/(18! * 12! * 12! * 8!)$
3	$300!/(5! * 8! * 287!)$
4	$(3!3!2!)/10! = 1/50400$
5	$M(n, (n_1, \dots, n_6))/6^n$
6	$(7!)/(2 * 6^7)$
7	$M(12, (6, 2, 4)) * M(13, (4, 6, 3))/M(25, (10, 8, 7))$
8	$M(12, (3, 3, 3, 3)) * M(40, (10, 10, 10, 10))/M(52, (13, 13, 13, 13))$
9	$4!/M(52, (13, 13, 13, 13))$
10	$(2! * 3! * 4!)/9!$

## 1.10 The Probability of a Union of Events

1	$\approx 0.11913$
2	85

### 1.10.1

$$Pr(A_1) = Pr(A_2) = Pr(A_3) = C(4, 2) * C(48, 3)/C(52, 5)$$

$$Pr(A_1 \cup A_2) = Pr(A_1 \cup A_3) = Pr(A_2 \cup A_3) = C(4, 2) * C(48, 3) * C(45, 3)/C(52, 5)^2$$

$$Pr(A_1 \cup A_2 \cup A_3) = 0$$

$$Pr(A_1 \cup A_2 \cup A_3) = 3 * C(4, 2) * C(49, 3)/C(52, 5) - 3C(4, 2) * C(49, 3) * C(46, 3)/C(52, 5)^2$$

TODO later (probably never).