Statistical methods for archaeological data analysis I: Basic methods

06 - Basic Probability Theory

Martin Hinz

Institut für Archäologische Wissenschaften, Universität Bern

8.04.2025

Repetition: Population and sample [1]

Population

Amount of all items of relevance for an analysis.

Sample

Selection of items on basis of certain criteria (e.g. representativity) which will be analysed instead of the population

The difference should always be kept in mind

In archaeology only sampling is possible! The population can never be investigated!

Repetition: Population and sample [2]

Features of the population: *parameters*

Parameters always exist, they have a certain value, but they are unknown and often (most of the time) also uncheckable.

Example:

μ : mean of the population σ: standard deviation of the population

$\bar{x}$: mean of the sample s: standard deviation of the sample

In statistical tests only features of the sample could be checked. The quality of the statement of a test therefore depends on the choice of the sample (representativity)!

Repetition: Null hypothesis

Validation through falsification

In statistical tests most of the times not the statement is tested which one expects to be true but one tries to disprove the statement which one expects to be wrong: the null hypothesis.

This hypothesis states mostly, that a association do not exists or that there is no differences between the samples and the distribution of the observations is by chance.

Example: Is the composition of grave goods different between male and female deceased?

H_0: The composition is the same

H_1: The composition is different

Reason

1. It is (logical) easier to prove, that a statement is wrong (falsify) then to prove that a statement is true (verify).
2. Most of the times it is easier to formulate a null hypothesis (How exactly is the composition different?). It doesn't make a assumption about how the character of a association/difference exactly is.

concept of probability

Subjective (everyday) concept of probability

Known from everyday life, we use every day. Subjective assumptions about probabilities (not verifiable)

Example

- Probably there is still no snow outside if the lecture is is over.
- There will probably be a bad result if I don't do my homework

Statistical (mathematical, quasi-objective) concept of probability

Based on mathematical probability laws

Refers to (in principle) repeatable operations whose result (in the principle) is not predictable

Classic example: Gambling

Probability and probability theory

The basis of every statistic

Prediction of unknown quantities by known values with a certain probability of error for given framework parameters

Development of probability theory based on random experiments

Definition of Kolmogorov

Classical probability experiment: Rolling dice

(for role players: W6)

The result of a dice roll is a probability event (elementary event): 5

All possible events form the event space.

$\Omega=\{1,2,3,4,5,6\}$ => set

Multiple dice results also form a set

$A = \{2,4,1,3,5\}$

Set Theory

Some symbols...

Set A={1,2,3,4}; Set B={4,5,6}; Event space Ω={1,2,3,4,5,6}	
1 is an element of set A	$1 \in A$
C is the union of A and B {1,2,3,4,4,5,6}	$C = A \cup B$
D is the intersection of A and B {4}	$D = A \cap B$
E is A minus B {1,2,3}	$E = A - B$
Not A (=event space - A)	$\bar{A} = \Omega - A = 5, 6$
The intersection of D {4} and E {1,2,3} is the empty set	$D \cap E = \emptyset$

probability calculation [1]

Classical Probability Definition by Laplace

$$p(A) = \frac{Number\ of\ positive\ results}{Number\ of\ possible\ results}$$

Relative Frequency of an event

Dice example

A=6 , Event space={1,2,3,4,5,6}

$$p(6) = \tfrac{1}{6} = 0.1667 = 16,67\%$$

$$p(\bar{6}) = p(\Omega) - p(6) = 1 - \tfrac{1}{6} = \tfrac{5}{6} = 0.8333 = 83.33\%$$

probability calculation [2]

basic probabilities

The probability of anything happening is always 100%.

$$p(\Omega) = 1$$

Is an event safe: 100 % probability

$$p(A) = 1$$

$$p(this\ is\ a\ statistics\ course) = 1$$

If an event is excluded: 0% probability

$$p(A) = 0$$

$$p(here\ you\ can\ learn\ something\ about\ knitting) = 0$$

probability calculation [3]

complementary events

Roll the dice:

Without physical tricks a dice roll always has a number as result

$$p(6) = \tfrac{1}{6} \rightarrow p(A) = \tfrac{1}{6}$$

$$p(1...5) = \tfrac{5}{6} \rightarrow p(\bar{A}) = \tfrac{5}{6}$$

$$p(A) + p(\bar{A}) = 1$$

The probability of an event and its opposite is always 1, so you can calculate one from the other.

Example: A card game has 4 colors (diamonds, hearts, spades, clubs).

The probability to draw a heart card is 1 of 4: 0.25

The probability of not drawing a Heart card is 3 of 4: 0.75, or 1 - 1 of 4: 1-0.25 = 0.75

Kolmogorov probability axioms

1. axiom

Each event from the event space is assigned a number p(A), which describes the probability of the event. This is between 0 and 1. $0 \leq p(A) \leq 1$

2. axiom

The safe event has the value one. $p(E) = 1$

3. axiom

For pairwise disjunctive events, i.e. those that do not have an intersection (e.g. {1,2} and {3,4}), the probability for their union is the sum of their individual probabilities.
$p(A_1 \cup A_2 \ldots \cup A_n) = \sum_{i=1}^{n} p(A_i)$

also z.B. $\Omega = \{1, 2, 3, 4, 5, 6\}, A = \{1, 2\}, B = \{3, 4\}$

$p(A) = \frac{2}{6}, P(B) = \frac{2}{6}, p(A \cup B) = p(A) + p(B) = \frac{2}{6} + \frac{2}{6} = \frac{4}{6} = 66,67\%$

Conditional and independent events

Conditional and independent events

Example Rolling dice: The result of the 2nd roll is independent of the result of the 1st roll. Therefore the probability is to roll first a 5 and then a 6:

$$p(A \cap B) = p(A) * p(B) = \frac{1}{6} * \frac{1}{6} = \frac{1}{36}$$

Example biscuits (after Dolić): A (non-transparent) bag with a chocolate biscuit, a sugar biscuit, an eco biscuit. How likely is it to get out the chocolate biscuit first and then the eco biscuit?

Wrong would be: $p(choco\ than\ eco) = p(choco) * p(eco) = \frac{1}{3} * \frac{1}{3} = \frac{1}{9}$

because after the chocolate biscuit's out, there's only two biscuits left $p(eco\ if\ choco) = \frac{1}{2}$

That's why: $p(choco\ than\ eco) = p(choco) * p(eco\ if\ choco) = \frac{1}{3} * \frac{1}{2} = \frac{1}{6}$

axiom of (conditional) probability
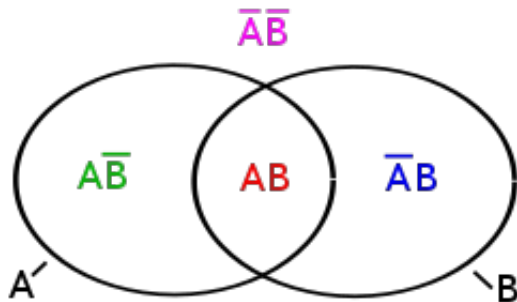
$$p(A \cap B) = p(A) * p(B|A)$$

addition law of probability

or the sum rule

Derived from the axioms:

For all possible combination of events

$$p(A \cup B) = p(A) + p(B) - p(A \cap B)$$

Example: Drawing of cards, how likely is it to draw a hearts card? 32 cards, 1/4 is hearts (8)

addition law of probability

or the sum rule

Derived from the axioms:

For all possible combination of events

$$p(A \cup B) = p(A) + p(B) - p(A \cap B)$$

Example: Drawing of cards, how likely is it to draw a hearts card? 32 cards, 1/4 is hearts (8)

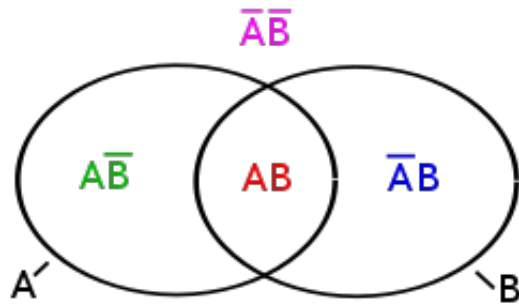$$p(A) = \frac{8}{32}$$

$$p(A) = \frac{1}{4}$$

addition law of probability

or the sum rule

Derived from the axioms:

For all possible combination of events

$$p(A \cup B) = p(A) + p(B) - p(A \cap B)$$

Example: Drawing of cards, how likely is it to draw a queen? 32 cards, 4 queen
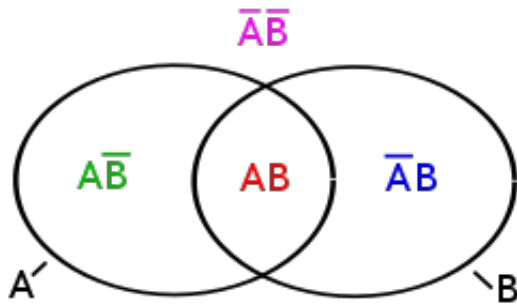
addition law of probability

or the sum rule

Derived from the axioms:

For all possible combination of events

$$p(A \cup B) = p(A) + p(B) - p(A \cap B)$$

Example: Drawing of cards, how likely is it to draw a queen? 32 cards, 4 queen

$$p(A) = \frac{4}{32}$$

$$p(A) = \frac{1}{8}$$

addition law of probability

or the sum rule

Derived from the axioms:

For all possible combination of events

$$p(A \cup B) = p(A) + p(B) - p(A \cap B)$$

Example: Drawing of cards, how likely is it to draw a trump card (queens and hearts are trump)? 32 cards, 4 queen, 8 heart, one queen of hearts
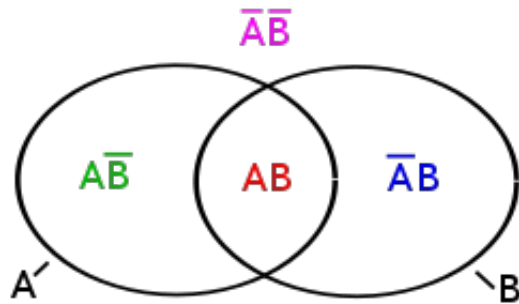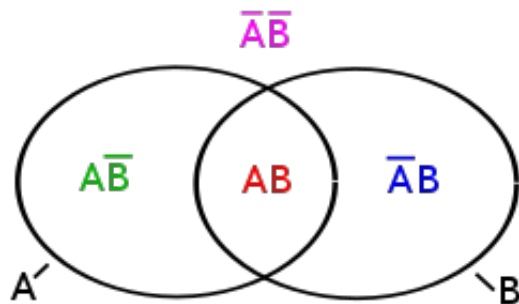
addition law of probability

or the sum rule

Derived from the axioms:

For all possible combination of events

$$p(A \cup B) = p(A) + p(B) - p(A \cap B)$$

Example: Drawing of cards, how likely is it to draw a trump card (queens and hearts are trump)? 32 cards, 8 heart, 4 queen, one queen of hearts
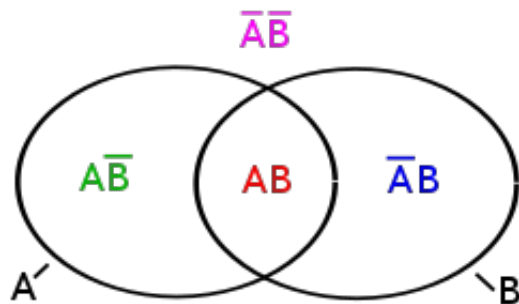
$$p(A) = \tfrac{1}{4}; p(B) = \tfrac{1}{8}; p(A \cup B) = p(A) + p(B) - p(A \cap B)$$

$$p(A \cup B) = \tfrac{1}{4} + \tfrac{1}{8} - \tfrac{1}{32} = \tfrac{11}{32} = 0.34375$$

Summary

Classical Probability Definition by Laplace

$$p(A) = \frac{Number\ of\ positive\ results}{Number\ of\ possible\ results}$$

Kolmogorov 1. axiom

Each event from the event space is assigned a number p(A), which describes the probability of the event. This is between 0 and 1. $0 \le p(A) \le 1$

Kolmogorov 2. axiom

The safe event has the value one. $p(E) = 1$

Kolmogorov 3. axiom

For pairwise disjunctive events, i.e. those that do not have an intersection (e.g. {1,2} and {3,4}), the probability for their union is the sum of their individual probabilities. $p(A_1 \cup A_2 \ldots \cup A_n) = \sum_{i=1}^{n} p(A_i)$

Multiple indepent events

$$p(A \cap B) = p(A) * p(B)$$

multiple conditional events

$$p(A \cap B) = p(A) * p(B|A)$$

multiple not disjunctive events

$$p(A \cup B) = p(A) + p(B) - p(A \cap B)$$
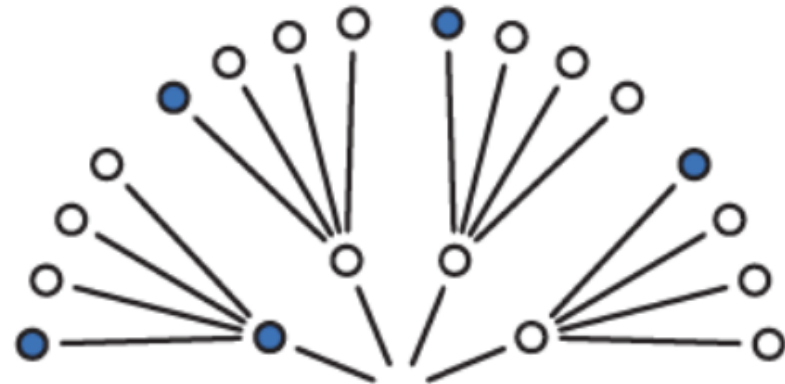
Source: McElreath 2017

combinatorics [1]

To determine the probability, it is necessary to know all possible events

This is easy for an event alone, but what if events are to be combined? -> Combinatorics

How many possibilities are there to select k elements from n elements?

	Variation (with respect to order)	Combination (without respecting the order)
with 'putting back'; with replacement	n^k	$\frac{(n+k-1)!}{k!*(n-1)!}$
without 'putting back'; without replacement	$\frac{n!}{(n-k)!}$	$\frac{n!}{k!*(n-k)!} = \binom{n}{k}$

combinatorics [2]

Examples

How many possibilities are there to combine 2 dice results?

With putting back, with respect to order

$$\Omega = \{1, 2, 3, 4, 5, 6\}; n(\Omega) = 6; number\ dices\ k = 2$$

$$B_n^{k=2} = \{(x_1, x_2) | x_i \in \Omega\}$$

$$n(B) = n(x_1) * n(x_2) = n(\Omega) * n(\Omega) = 6 * 6 = 36$$

Therefore: Probability for 1. roll=6, 2. roll=6: $p(6, 6) = \frac{1}{6^2} = \frac{1}{36} = 0,0278 = 2,78\%$

combinatorics [3]

How many possible unique Lotto tickets (6 of 49) are there?

Without 'putting back', without respecting order

$$\Omega = \{1, 2, 3..., 49\}, n(\Omega) = 49, number\ balls\ k = 6$$

$$B_n^{k=6} = \{(x_1, x_2, \ldots, x_6) || x_i \in \Omega | (x_1, \ldots x_{i-1})\}$$

$$n(B) = \frac{n!}{k! * (n-k)!} = \frac{49!}{6! * (49-6)!} = 13983816$$

Therefore: Probability for 6 correct crosses:

$$p(6_{right}) = \frac{1}{13983816} = 0.000000072 = 0.000007151\%$$

3 / 52

Law of large numbers

The larger the sample, the more similar the distribution of sample and population

Example: dice

The theoretical probability for each result is 1/6, the total population of all rolls ever made (with unbiased dice), the proportion of each number should be pretty much 1/6.

The more often one throws the dice experimentally, the more similar is the distribution of the sample to the population.

The relative frequency of the random results converges against the probability of the random result

The law of the large number is the *bridge from the sample to the population*, it allows statements to be made about the population without knowing it.

Simulating law of large numbers

```
rolls=1000

roll<- as.numeric(sample(1:6,rolls, replace=T)

list=0

count=0

position=0

anteil=0

for (test in roll)
{
list<-append(list,test,length(list))
anteil<-append(anteil,sum(list==1)/length(list
}
plot(anteil,type="l",ylim=c(0,1))
abline(h=1/6,lty=3)
```


Simulated dice experiment, the proportion of the number of 6 eyes is plotted, the dotted line shows the probability for 6 eyes.
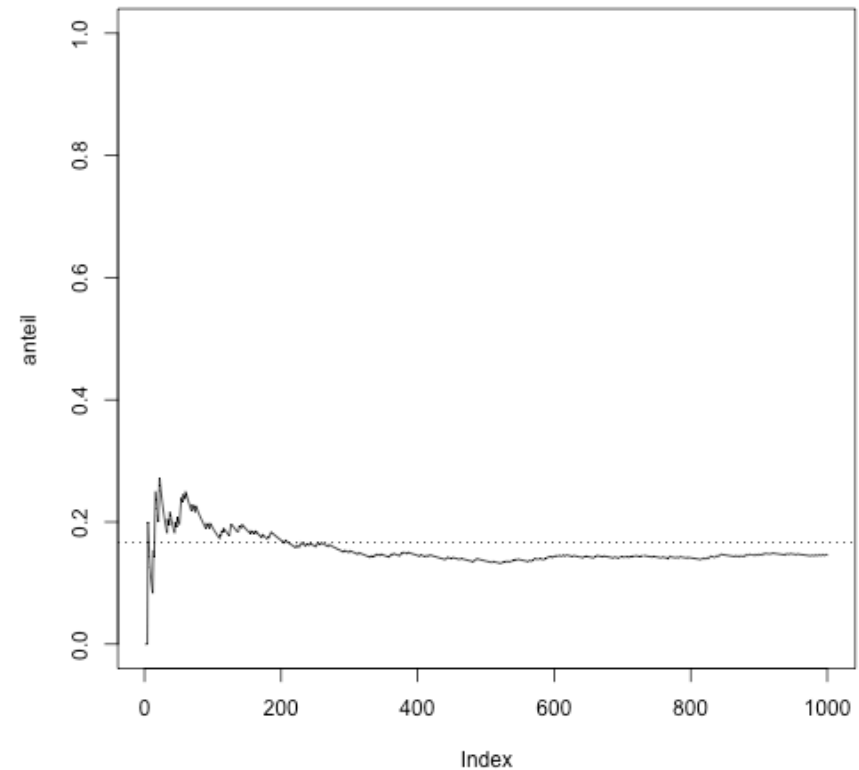
Random variables

What is random at all?

A random variable is the result of a (complex?, unknown?) process.

The various possible outcomes of the process represent the values of the random variable.

Whether a variable is regarded as random or not depends on the definition:

Coin toss: The result of a coin toss is *determined* by different physical laws (throwing force, density of air, gravity etc.)! Since we cannot control these, the result to be considered *random*!

In order to be able to analyze them statistically, the results are converted into the space of real numbers (recoded).

Random Variable Example

Example (after Dolić)

A coin is flipped three times. The number of "heads" (H) is noted as a random variable. Possible results:

coin flip	x_i	$p(x_i)$
TTT	0	1/8
TTH	1	1/8
THT	1	1/8
HTT	1	1/8
THH	2	1/8
HTH	2	1/8
HHT	2	1/8
HHH	3	1/8

Probability function (density function)

Example tossing a coin

typical properties

Expected value: The value that is most probable.

Dispersion: The variance of distribution

more: Skewness and Kurtosis

$$f(x_i) = \begin{cases} p(x_i = 0) = \frac{1}{8} \\ p(x_i = 1) = \frac{3}{8} \\ p(x_i = 2) = \frac{3}{8} \\ p(x_i = 3) = \frac{1}{8} \end{cases}$$
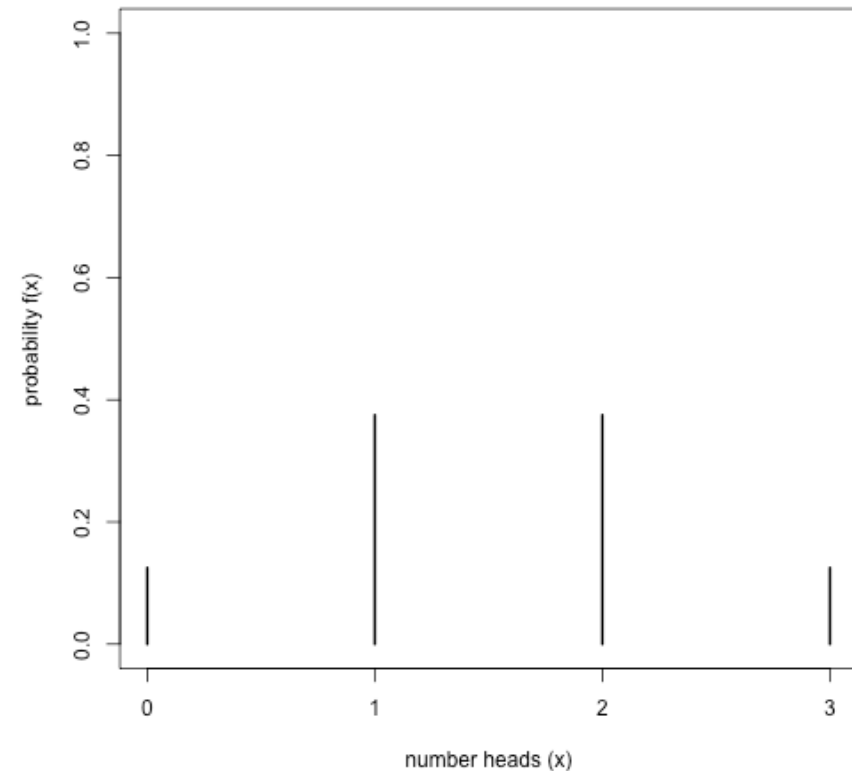
Cumulative distribution function

Is the sum function of the probability function

"What is the probability of having up to two heads?"

Properties:

$$0 \leq F(x) \leq 1$$

F(x) is monotonous not falling

$$F(x_1) \leq F(x_2)\ldots \leq F(x_n)$$

$$f(x_i) = \begin{cases} p(x_i < 0) = 0 \\ p(x_i \leq 0) = \frac{1}{8} \\ p(x_i \leq 1) = \frac{4}{8} \\ p(x_i \leq 2) = \frac{7}{8} \\ p(x_i \leq 3) = 1 \end{cases}$$
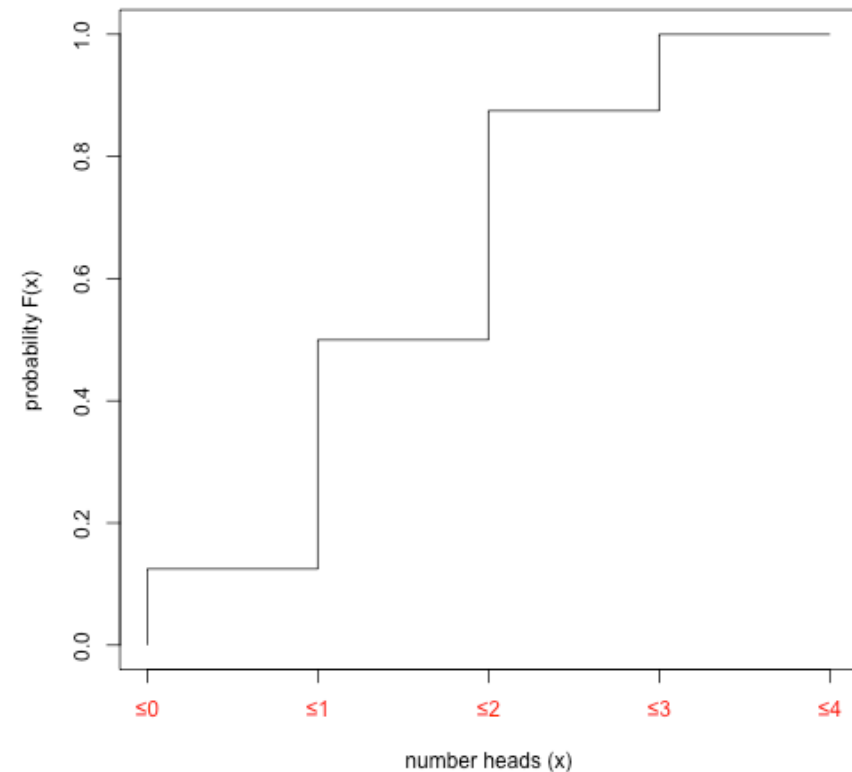
Relationship with statistical tests

Question: Is someone playing with biased coins?

How can one significantly (error probability 5%) determine that the coins are biased and always show head?

H_0: The coins are not biased, the distribution corresponds to the distribution of an unbiased coin toss (Binomial distribution).

H_1: The coins are biased, the distribution differs significantly from the distribution of an unbiased coin toss (Binomial distribution).

N=20 throws

We need:

Rejection range: number head that is high enough to be sure with 95% of probability the coin is biased

Relationship with statistical tests

Question: Is someone playing with biased coins?

We need:

Rejection range: number head that is high enough to be sure with 95% of probability the coin is biased

If the probability of the random occurrence of a result is less than 5%, the occurrence of that event is not random with 95% probability.

Relationship with statistical tests

Question: Is someone playing with biased coins?

We need:

Rejection range: number head that is high enough to be sure with 95% of probability the coin is biased

If the probability of the random occurrence of a result is less than 5%, the occurrence of that event is not random with 95% probability.

So: For how many head results on 20 rolls is the probability 5% or lower?

Relationship with statistical tests

Question: Is someone playing with biased coins?

We need:

Rejection range: number head that is high enough to be sure with 95% of probability the coin is biased

If the probability of the random occurrence of a result is less than 5%, the occurrence of that event is not random with 95% probability.

So: For how many head results on 20 rolls is the probability 5% or lower?

Basically: Probability according to Laplace:

$$p(A) = \frac{Number\ of\ positive\ results}{Number\ of\ possible\ results}$$

Relative frequency of an event

Relationship with statistical tests

Question: Is someone playing with biased coins?

We need:

Rejection range: number head that is high enough to be sure with 95% of probability the coin is biased

Basically: Probability according to Laplace:

$$p(A) = \frac{Number\ of\ positive\ results}{Number\ of\ possible\ results}$$

We need: Number of positive events (heads) and number of possible events (total number of possible results of throws of 20 coins)

number of possible events

1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20
H	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T

1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20
T	H	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T

1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20
H	H	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T

1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20
T	T	H	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T

1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20
H	T	H	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T

1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20
H	H	H	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T

Relationship with statistical tests

Question: Is someone playing with biased coins?

We need:

Rejection range: number head that is high enough to be sure with 95% of probability the coin is biased

Number of possible events (total number of possible results of throws of 20 coins)

-> combinatorics:

Is the order important?

Relationship with statistical tests

Question: Is someone playing with biased coins?

We need:

Rejection range: number head that is high enough to be sure with 95% of probability the coin is biased

Number of possible events (total number of possible results of throws of 20 coins)

-> combinatorics:

Is the order important?

Answer: **Yes** (suprised?)

[T,T,H] represents a different case that [H,T,T]

If a coin shows tails, does that change the probability of the next coin to show tails (without replacement) or does this probability remain unchanged (with replacement)?

Relationship with statistical tests

Question: Is someone playing with biased coins?

We need:

Rejection range: number head that is high enough to be sure with 95% of probability the coin is biased

Number of possible events (total number of possible results of throws of 20 coins)

-> combinatorics:

Is the order important?

Answer: **Yes** (suprised?)

[T,T,H] represents a different case that [H,T,T]

If a coin shows tails, does that change the probability of the next coin to show tails (without replacement) or does this probability remain unchanged (with replacement)?

Answer is: **No**. Chances does not change, so with replacement.

Relationship with statistical tests

Question: Is someone playing with biased coins?

We need:

Rejection range: number head that is high enough to be sure with 95% of probability the coin is biased

Number of possible events (total number of possible results of throws of 20 coins)

-> combinatorics:

with respect to order and with replacement:

	Variation (with respect to order)	Combination (without respecting the order)
with 'putting back'; with replacement	n^k	$\frac{(n+k-1)!}{k!*(n-1)!}$
without 'putting back'; without replacement	$\frac{n!}{(n-k)!}$	$\frac{n!}{k!*(n-k)!} = \binom{n}{k}$

Relationship with statistical tests

Question: Is someone playing with biased coins?

n^k

2 possible cases (Heads, Tails) : n 20 possible positions : k

number of possible results: $n^k = 2^{20}$ **= 1048576**

2 results can be distributed in 1048576 ways on 20 positions

number of positive events

1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20
H	H	H	H	H	H	T	T	T	T	T	T	T	T	T	T	T	T	T	T

1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20
H	H	H	H	H	T	H	T	T	T	T	T	T	T	T	T	T	T	T	T

1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20
H	H	H	H	H	T	T	T	T	T	T	T	T	T	T	T	T	T	T	H

Relationship with statistical tests

Question: Is someone playing with biased coins?

We need:

Rejection range: number head that is high enough to be sure with 95% of probability the coin is biased

Number of positive events: How many possibilities are there to divide a fixed number of coins with head (k: 1-20) into 20 places (n)?

n=20 places, k= cases head

Is it important in which order the coins fall?

Relationship with statistical tests

Question: Is someone playing with biased coins?

We need:

Rejection range: number head that is high enough to be sure with 95% of probability the coin is biased

Number of positive events: How many possibilities are there to divide a fixed number of coins with head (k: 1-20) into 20 places (n)?

n=20 places, k= cases head

Is it important in which order the coins fall?

Answer: No, we are only interested in the number of e.g. 1x head, 2x tails, not in their order.

If a coin shows tails, does this change the distribution of head/tail to the remaining places? (yes: without, no: with replacement)

Relationship with statistical tests

Question: Is someone playing with biased coins?

We need:

Rejection range: number head that is high enough to be sure with 95% of probability the coin is biased

Number of positive events: How many possibilities are there to divide a fixed number of coins with head (k: 1-20) into 20 places (n)?

Is it important in which order the coins fall? No

Answer: No, we are only interested in the number of e.g. 1x head, 2x tails, not in their order.

If a coin shows tails, does this change the distribution of head/tail to the remaining places? (yes: without, no: with replacement)

Answer: Yes, the distribution is changing! If I have chosen head for a place, the number of remaining heads changes (with a fixed number of heads)!

Relationship with statistical tests

Question: Is someone playing with biased coins?

n = 20 places, k = number heads

Number of positive events

-> combinatorics:

with not respecting the order and without replacement:

	Variation (with respect to order)	Combination (without respecting the order)
with 'putting back'; with replacement	n^k	$\frac{(n+k-1)!}{k!*(n-1)!}$
without 'putting back'; without replacement	$\frac{n!}{(n-k)!}$	$\frac{n!}{k!*(n-k)!} = \binom{n}{k}$

Relationship with statistical tests

Question: Is someone playing with biased coins?

n = 20 places, k = number heads

Number of possible events = $n^k = 2^{20} = 1048576$

Number of positive events: calculated according to binomial coefficient, possibilities to arrange number of outputs in number of throws

No times head: n = 20, k=0: only one possibility $\qquad \frac{n!}{k!*(n-k)!} = \binom{n}{k}$

1 times head: n = 20, k = 1: 20 possibilities

2 times head: n = 20, k = 2: 190 possibilities

3 times head: n = 20, k = 3: 1140 possibilities

4 times head: n = 20, k = 4: 4845 possibilities

5 times head: n = 20, k = 5: 15504 possibilities

6 times head: n = 20, k = 6: 38760 possibilities

Relationship with statistical tests

Question: Is someone playing with biased coins?

Number of possible events = $n^k = 2^{20} = 1048576$

$$p(A) = \frac{Number\ of\ positive\ results}{Number\ of\ possible\ results}$$

number head	possibilities	positive/possible	cumulative
0	1	0.000	0.000
1	20	0.000	0.000
2	190	0.000	0.000
3	1140	0.001	0.001
4	4845	0.005	0.006
5	15504	0.015	0.021
6	38760	0.037	0.058

Relationship with statistical tests

Question: Is someone playing with biased coins?

Number of possible events = $n^k = 2^{20} = 1048576$

General: $B_{n;k;p} = \binom{n}{k} * p^k * (1-p)^{n-k}$: Equation for binomial distribution

In this example eg. 2 Heads, 18 Tails, k = 2, n = 20, p = 0.5:

190/1048576 = 0.000181 = $\binom{20}{2} * 0.5^2 * (1-0.5)^{20-2} = 190 * 0.25 * 3.815 \times 10^{-6}$

number head	possibilities	positive/possible	cumulative
0	1	0.000	0.000
1	20	0.000	0.000
2	190	0.000	0.000
3	1140	0.001	0.001
4	4845	0.005	0.006
5	15504	0.015	0.021
6	38760	0.037	0.058

Relationship with statistical tests

Question: Is someone playing with biased coins?

N=20 throws

We need:

Rejection range: Number Tails < 6: 95% Probability, that something with the coins is wrong. (25% of the tosses)

In R:

```
binom.test(5,20,0.5)
```

```
##
##      Exact binomial test
##
## data:  5 and 20
## number of successes = 5, number of trials = 20, p-value = 0.04
## alternative hypothesis: true probability of success is not equal to 0.5
## 95 percent confidence interval:
##  0.0866 0.4910
## sample estimates:
## probability of success
##                   0.25
```

Relationship with statistical tests

Question: Is someone playing with biased coins?

N=200 throws

We need:

Rejection range: Number Tails < 85: 95% Probability, that something with the coins is wrong. (42% of the tosses)

In R:

```
binom.test(85,200,0.5)
```

```
##
##      Exact binomial test
##
## data:  85 and 200
## number of successes = 85, number of trials = 200, p-value = 0.04
## alternative hypothesis: true probability of success is not equal to 0.5
## 95 percent confidence interval:
##   0.356 0.497
## sample estimates:
## probability of success
##                  0.425
```