Elements of Statistics

Chapter 5: Probability theory

Ralf Münnich, Jan Pablo Burgard and Florian Ertz

University of Trier
Faculty IV
Economic and Social Statistics Department

Winter term 2021/22

Notes

Example 5.1: Some experiments

a) Rolling a dice
b) Playing the lottery
c) Duration between two computer malfunctions
d) Kilometer reading of a car
e) Burning life of a bulb
f) Turnover of a pharmacy on a Friday
g) Life span of a male live birth
h) Time to solve a problem

- What are the outcomes of these experiments?
- Are these experiments reproducible?

Notes

Experiment and sample space

Definition 5.1: Experiment

Procedures, which can actually or at least theoretically be repeated under a constant set of conditions and the outcomes of which cannot be forecast precisely, are called experiments.

Definition 5.2: Sample space

The set of all possible, mutually exclusive outcomes of an experiment is called sample space Ω.

Sample spaces may be:

- Finite
- Infinite
  - (Un)countable

Notes

Example 5.2:

One roll of a dice:

a) $\Omega = \{i \,|\, i \in \mathbb{N};\quad 1 \leq i \leq 6\}$
   i is number of pips
b) $\Omega = \{\omega_g, \omega_u\}$
   Even / uneven number of pips
c) $\Omega = \{\{1,2,3\},\{4,5\},\{6\}\}$

Example 5.3:

Burning life of a bulb:

a) $\Omega = \{x | x \in \mathbb{R}; 0 \leq x \leq 10{,}000\}$
   Infinite accuracy of measurement
   $\rightarrow$ continuous variable
b) $\Omega = \{x | x \in \mathbb{N}_0, x \leq 10{,}000\}$
   Measurement in hours
   $\rightarrow$ discrete variable

Notes

Wirtschafts- und Sozialstatistik
Universität Trier

Definition 5.3:

Each subset of a sample space is called event.

We write:

$A, B, C \ldots$

We say:

An experiment with sample space Ω yielded event A (B, C, ...).

Example 5.4:

One roll of a dice $\Omega = \{1, 2, 3, 4, 5, 6\}$:

$$A = \{1, 3, 5\} \qquad B = \{2, 4, 6\} \qquad C = \{3\}$$

Application in R:

```
Omega <- 1:6

A <- Omega[c(1,3,5)] ; B <- Omega[c(2,4,6)] ; C <- Omega[3]
```

Bear in mind: Set theory and its calculation rules

Notes

Wirtschafts- und Sozialstatistik
Universität Trier

Example 5.5: see Ex. 5.4 (1)

$B = \overline{A}$

```
A_Bar <- Omega[-A]
setequal(x = B, y = A_Bar)
[1] TRUE
```

$\overline{C} = \{1, 2, 4, 5, 6\}$

```
C_Bar <- Omega[-C]
C_Bar
[1] 1 2 4 5 6
```


Notes

Wirtschafts- und Sozialstatistik
Universität Trier

Example 5.5: see Ex. 5.4 (2)

$A \cap B = \emptyset$

```
intersect(x = A, y = B)
integer(0)
```

$B \cup C = \{2, 3, 4, 6\}$

```
sort(union(x = B, y = C))
[1] 2 3 4 6
```

Events may be characterised using *admissible questions*.

$\rightarrow$ System of events (*set of admissible questions*)

Notes

Wirtschafts- und Sozialstatistik
Universität Trier

Event space

Definition 5.4:

Let Ω be a sample space and let $\mathcal{K}$ be a non-empty subset of the power set of the sample space $\mathcal{P}(\Omega)$. $\mathcal{K}$ is called an event space if the following properties hold:

I: $\emptyset \in \mathcal{K}$

II: If $A \in \mathcal{K}$, then $\overline{A} \in \mathcal{K}$

III: If $A \in \mathcal{K}$ and $B \in \mathcal{K}$, then $A \cup B \in \mathcal{K}$

We may call $\mathcal{K}$ an algebra over Ω as well.

Example 6.6:

$\mathcal{K}_1 = \{\emptyset, \{1, 3, 5\}, \{2, 4, 6\}, \Omega\}$
is an event space.

$\mathcal{K}_2 = \{\emptyset, \{1, 2\}, \{3, 4\}, \{5, 6\}, \Omega\}$
is not an event space.

$\mathcal{P}(\Omega)$ is always an event space (containing 2^n elements).

Notes

Sigma algebra

Wirtschafts- und Sozialstatistik
Universität Trier

Definition 5.5:
Let Ω be a sample space. A non-empty subset $\mathcal{S}$ of $\mathcal{P}(\Omega)$ is called sigma algebra if the following holds:

I: $A \in \mathcal{S} \Rightarrow \overline{A} \in \mathcal{S}$ (closed under complementation)

II: $A_i \in \mathcal{S}$ for all $i = 1, 2, \ldots$

$$\Rightarrow \bigcup_{i=1}^{\infty} A_i \in \mathcal{S} \text{ (closed under countable union)}$$

- Every countable union of events itself is part of $\mathcal{S}$.
- All set operations can be derived from these two set operations

Notes

Definition 5.6:

Wirtschafts- und Sozialstatistik
Universität Trier

Let $\mathcal{K}$ be an event space over Ω. The (set) function $P : \mathcal{K} \to \mathbb{R}$ is called a *probability content*, if the following holds:

1. If $A \in \mathcal{K}$, then $P(A) \geq 0$
2. If $A, B \in \mathcal{K}$ and $A \cap B = \emptyset$, then
$$P(A \cup B) = P(A) + P(B)$$
3. $P(\Omega) = 1$

Notes

Implications

Wirtschafts- und Sozialstatistik
Universität Trier

1. $P(\overline{A}) = 1 - P(A)$
2. $P(A \cap B)$ follows from Definition 6.6.
3. Finite unions follow from Definition 6.6.
4. Countable unions do not follow from Definition 6.6
(e. g. even numbers within the natural numbers).

Notes

Definition 5.7:

Wirtschafts- und Sozialstatistik
Universität Trier

Let $\mathcal{S}$ be a sigma algebra over a sample space Ω. A (set) function $P : \mathcal{S} \to \mathbb{R}$ is called *probability measure*, if the following holds:

1. If $A \in \mathcal{S}$, then $P(A) \geq 0$ (**Non-negativity**)
2. If $A_i \in \mathcal{S}$, $i = 1, 2, \ldots$ and all A_i are pairwise disjoint, we have
$$P\Big(\bigcup_{i=1}^{\infty} A_i\Big) = \sum_{i=1}^{\infty} P(A_i)$$
(σ-**additivity**)
3. $P(\Omega) = 1$ (**Standardisation**)

Bear in mind:
Pairwise disjoint and *overall disjoint* have to be distinguished!

Notes

Probability measure and implications

Wirtschafts- und Sozialstatistik Universität Trier

Definition 5.8:

A triple $(\Omega; \mathcal{S}; P)$ consisting of a sample space Ω, a sigma algebra $\mathcal{S}$ over Ω and a probability measure P over $\mathcal{S}$ is called *probability space*.

Further implications

1. $P(A \cup B) = P(A) + P(B) - P(A \cap B)$
2. $P(A) \leq 1$
3. $P(\emptyset) = 0$
4. $A \subset B \quad \Rightarrow \quad P(A) \leq P(B)$
5. $P(A \cup B \cup C) = P(A) + P(B) + P(C) - P(A \cap B) - P(A \cap C) - P(B \cap C) + P(A \cap B \cap C)$

Notes

Classical or Laplace's concept - Principle of symmetry

Wirtschafts- und Sozialstatistik Universität Trier

- Finite sample space Ω
- Elementary events (outcomes) have equal probability of occurrence

$\nu(\Omega)$ is the number of elementary events and $\nu(A)$ is the number of *favourable* cases.

Then

$$P(A) = \frac{\nu(A)}{\nu(\Omega)}$$

is the probability that event A occurs (number of favourable cases divided by number of possible cases).

Notes

Example 5.7: Two rolls of a dice (1)

Wirtschafts- und Sozialstatistik Universität Trier

Two rolls of a dice: $\Omega = \{(i,j) | i,j \in \mathbb{N} \quad \text{mit} \quad 1 \leq i,j \leq 6\}$

There are 36 possible events.

Application in R:

```
Omega <- expand.grid(Dice1 = 1:6, Dice2 = 1:6)
head(Omega, n = 9)
```

```
  Dice1 Dice2
1     1     1
2     2     1
3     3     1
4     4     1
5     5     1
6     6     1
7     1     2
8     2     2
9     3     2
```

```
length(Omega[,1])
```

```
[1] 36
```


Notes

Example 5.7: Two rolls of a dice (2)

Wirtschafts- und Sozialstatistik Universität Trier

a) Sum of pips is **at least** 10:

$A = \{(4;6);(5;5);(5;6);(6;4);(6;5);(6;6)\}$

$W(A) = \frac{\nu(A)}{36} = \frac{6}{36} = \frac{1}{6}$

Calculations in R:

```
Sum_of_pips <- apply(Omega, 1, sum)
Omega <- cbind(Omega, Sum_of_pips)
A <- Omega[Omega$Sum_of_pips >= 10, ]
A
```

```
   Dice1 Dice2 Sum_of_pips
24     6     4          10
29     5     5          10
30     6     5          11
34     4     6          10
35     5     6          11
36     6     6          12
```

```
length(A[,1])/length(Omega[,1])
```

```
[1] 0.1666667
```


Notes

Example 5.7: Two rolls of a dice (3)

Wirtschafts- und Sozialstatistik
Universität Trier

b) Sum of pips is **exactly** 4:

$B = \{(1;3);(2;2);(3;1)\}$

$W(B) = \frac{3}{36} = \frac{1}{12}$

Calculations in R:

```
B <- Omega[Omega$Sum_of_pips == 4, ]
B
```

```
     Dice1   Dice2 Sum_of_pips
3        3       1           4
8        2       2           4
13       1       3           4
```

```
length(B[,1])/length(Omega[,1])
```

```
[1] 0.08333333
```


Notes

Statistical concept of probability - Principle of frequency

Wirtschafts- und Sozialstatistik
Universität Trier

- Experiments can be arbitrarily repeated
- A as an event in the experiment
- n repetitions (independent trials)

$p_n(A)$ is the relative frequency of occurrences of event A in n repeated experiments.

Properties: Non-negativity, additivity and standardisation!

Then $P(A) = \lim_{n\to\infty} p_n(A)$.

Law of large numbers (see later).

Notes

Subjectivistic concept of probability

Wirtschafts- und Sozialstatistik
Universität Trier

- Probability of rain
- Investment in shares
- Risk of an accident for a certain new car

Subjectivistic determination:

- Expert knowledge
- Experience
- Intuition

Verifiability is a problem here.

Notes

Example 5.8: Probability of survival

Wirtschafts- und Sozialstatistik
Universität Trier

We determine the *probability* for a 50 year old man to survive the next year using the life table 2008/2010.

$p_{50} = 0.995868$.

Contrary to this: $\frac{l_{51}}{l_0} = 0.95169$.

Segmentation possible regarding further information:

- Person smokes
- Person has not been ill since 20 years
- Person has had a heart attack
- ...

Notes

Example 5.9

Wirtschafts- und Sozialstatistik Universität Trier

We want to determine the probability of obtaining a 6 in a roll of a dice.

We are informed that the experiment yielded an even number of pips. Does that change the probability?

We have: $P(\{6\}|\{2,4,6\}) = \dfrac{P(\{6\})}{P(\{2,4,6\})} = \dfrac{\frac{1}{6}}{\frac{1}{2}} = \dfrac{1}{3}$.

Calculations in R:

```
Prob <- length(6)/length(c(2,4,6))
Prob
```

```
[1] 0.3333333
```


Notes

Definition 5.9:

Wirtschafts- und Sozialstatistik Universität Trier

Let A and B be two events and let $P(B) \neq 0$. Then

$$P(A|B) = \frac{P(A \cap B)}{P(B)}$$

is called the probability of A conditional on the occurrence of event B.

Analogously: $P(B|A)$ with $P(A) \neq 0$.

The multiplication theorem follows:

$$P(A \cap B) = P(B) \cdot P(A|B) = P(A) \cdot P(B|A)$$

Notes

Example 5.10: Evaluation of a course (1)

Wirtschafts- und Sozialstatistik Universität Trier

The evaluation of a course yielded the following frequency table:

	C male	D female	$\sum$
A good	0.45	0.35	0.8
B bad	0.15	0.05	0.2
$\sum$	0.6	0.4	1

Calculations in R:

```
load("Example5-10.RData")
p_j_k <- addmargins(p_j_k)
p_j_k
```

```
       C    D Sum
A   0.45 0.35 0.8
B   0.15 0.05 0.2
Sum 0.60 0.40 1.0
```


Notes

Example 5.10: Evaluation of a course (2)

Wirtschafts- und Sozialstatistik Universität Trier

Then:

$$P(A) = 0.8$$
$$P(A|C) = \frac{P(A \cap C)}{P(C)} = \frac{0.45}{0.6} = 0.75$$
$$P(A|D) = \frac{0.35}{0.4} = \frac{7}{8} = 0.875$$

Calculations in R:

```
Prob_A   <- p_j_k[1,3]
Prob_A_C <- p_j_k[1,1]/p_j_k[3,1]
Prob_A_D <- p_j_k[1,2]/p_j_k[3,2]
```

```
Prob_A          Prob_A_C          Prob_A_D
[1] 0.8         [1] 0.75          [1] 0.875
```


Notes

Definition 5.10:

Wirtschafts- und Sozialstatistik Universität Trier

Let A and B be two random events. A and B are called stochastically independent if and only if:

$$P(A \cap B) = P(A) \cdot P(B)$$

Then we have:

$$P(A|B) = \frac{P(A \cap B)}{P(B)} \overset{\text{st. ind.}}{=} \frac{P(A) \cdot P(B)}{P(B)} = P(A)$$

Notes

Example 5.11:

Wirtschafts- und Sozialstatistik Universität Trier

A coin is tossed twice (heads or tails). Consider the following two events:

A: 1st toss yields heads

B: 2nd toss yields tails

Then:

$$P(A \cap B) = P(A) \cdot P(B) = \frac{1}{2} \cdot \frac{1}{2} = \frac{1}{4}$$

Notes

Example 5.12: see Ex. 5.10

Wirtschafts- und Sozialstatistik Universität Trier

We want to check if rating and gender are stochastically independent. Inter alia we should have:

$$P(A \cap C) = P(A) \cdot P(C).$$

But we actually have:

$$P(A) \cdot P(C) = 0.8 \cdot 0.6 = 0.48 \neq 0.45 = P(A \cap C)\,.$$

It follows that A and C are stochastically dependent..

Calculations in R:

```
Prob_C <- p_j_k[3,1]
Prob_A * Prob_C == 0.45
```

```
[1] FALSE
```

Addendum: What would the joint probabilities be, if the variables were stochastically independent and the marginal distributions were unchanged?

Notes

Example 5.13: Doorknobs (1)

Wirtschafts- und Sozialstatistik Universität Trier

A supplier to the automobile industry produces 10,000 door handles per day on 4 different machines. Production is distributed as follows:

M_1 1000 pieces with 8% scrap,
M_2 2000 pieces with 5% scrap,
M_3 3000 pieces with 3% scrap,
M_4 4000 pieces with 2% scrap.

One door handle is randomly chosen from the daily production. What is the probability that the item is defective?

Data input in R:

```
Doorknobs_per_machine <- c(1000,2000,3000,4000)
Number_doorknobs
Rejects_per_machine <- c(0.08,0.05,0.03,0.02)
```


Notes

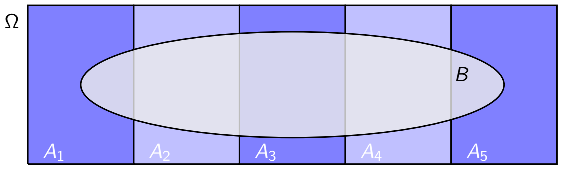
Theorem 5.1:

Wirtschafts- und Sozialstatistik Universität Trier

Law of total probability:

Let $A_1, \dots, A_m$ be a disjoint decomposition of Ω. Then, for $B \subset \Omega$ we have:

$$P(B) = \sum_{i=1}^{m} P(B \cap A_i) = \sum_{i=1}^{m} P(B|A_i) \cdot P(A_i)$$



Notes

Example 5.13: Doorknobs (2)

Wirtschafts- und Sozialstatistik Universität Trier

Let A_i for $i = 1, \dots, 4$ be the event that the door handle has been produced on machine M_i. Let F be the event that the door handle is faulty. Furthermore, we have:

$$P(A_1) = \frac{1000}{10000} = 0.1, \qquad P(F|A_1) = 0.08$$

$$P(A_2) = \frac{2000}{10000} = 0.2, \qquad P(F|A_2) = 0.05$$

$$P(A_3) = \frac{3000}{10000} = 0.3, \qquad P(F|A_3) = 0.03$$

$$P(A_4) = \frac{4000}{10000} = 0.4, \qquad P(F|A_4) = 0.02$$

Calculations in R:

```
P_Ai   <- Doorknobs_per_machine/Number_doorknobs
P_F_Ai <- Rejects_per_machine
```


Notes

Example 5.13: Doorknobs (3)

Wirtschafts- und Sozialstatistik Universität Trier

The probability for event F is composed as follows:

$$P(F) = P(A_1) \cdot P(F|A_1) + P(A_2) \cdot P(F|A_2) + P(A_3) \cdot P(F|A_3) + P(A_4) \cdot P(F|A_4) = \sum_{i=1}^{4} P(A_i) \cdot P(F|A_i).$$

Therefore, we have:
$P(F) = 0.1 \cdot 0.08 + 0.2 \cdot 0.05 + 0.3 \cdot 0.03 + 0.4 \cdot 0.02 = 0.035.$

Calculations in R:

```
P_F <- sum(P_Ai * P_F_Ai)
P_F
[1] 0.035
```


Notes

Example 5.14: Spam mails (1)

Wirtschafts- und Sozialstatistik Universität Trier

We want to discuss the nuisance of spam mails. Let the following two events be given:

A: The mail is spam.

B: The mail client marks the mail as spam.

Past experience taught us that roughly 85% of mails are spam. 95% of spam is marked as such by the mail client. But 8% of mails are wrongly marked as spam.

What percentage of mails, which have been marked as spam, is indeed spam?

Notes

Tree diagrams and probabilities

Wirtschafts- und Sozialstatistik
Universität Trier

Multiplication of probabilities
In a multi-stage experiment we get the probability of a single event by multiplying the probabilities on the respective path.

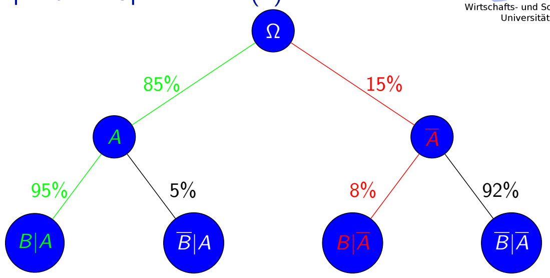
Addition of probabilities
Probabilities from different paths of a multi-stage experiment are added.

Total probability
The sum of probabilities at the *leaves* of a tree diagram is 1.

Notes

Example 5.14: Spam mails (2)

Wirtschafts- und Sozialstatistik
Universität Trier



We have: $P(A \cap B) = P(A) \cdot P(B|A)$
($P(\overline{A} \cap B)$, $P(A \cap \overline{B})$ and $P(\overline{A} \cap \overline{B})$ analogously). Overall we have

$$P(A|B) = \frac{0.85 \cdot 0.95}{0.85 \cdot 0.95 + 0.15 \cdot 0.08} = 0.985\,.$$

Notes

Theorem 5.2: Bayes' theorem

Wirtschafts- und Sozialstatistik
Universität Trier

Let $A_1, \ldots, A_m$ m be events of a sample space Ω. Furthermore, let the events A_i $(i = 1, \ldots, m)$ be pairwise disjoint and let

$$\Omega = \bigcup_{i=1}^{m} A_i. \qquad \text{(Decomposition of } \Omega\text{)}$$

Now, let B be another event and the probability for all positive events. Then we have for all $k = 1, \ldots, m$:

$$P(A_k|B) = \frac{P(A_k) \cdot P(B|A_k)}{\sum_{i=1}^{m} P(A_i) \cdot P(B|A_i)}\,.$$

We call $P(A_i)$ the a-priori-probability and $P(A_i|B)$ the a-posteriori-probability.

Notes

Example 5.15: see Ex. 5.13 (1)

Wirtschafts- und Sozialstatistik
Universität Trier

A quality control inspector of the supplier to the automobile industry has randomly sampled and inspected a doorknob where he detects a flaw. Now, he wants to know the probabilities for the flawed doorknob to be produced by machine M_1, M_2, M_3 or M_4. According to *Bayes' theorem* we obtain the respective probabilities as follows:

$$P(A_1|F) = \frac{P(A_1) \cdot P(F|A_1)}{\sum_{i=1}^{4} P(A_i) \cdot P(F|A_i)} = \frac{0.1 \cdot 0.08}{0.035} = 0.229$$

$$P(A_2|F) = \frac{P(A_2) \cdot P(F|A_2)}{\sum_{i=1}^{4} P(A_i) \cdot P(F|A_i)} = \frac{0.2 \cdot 0.05}{0.035} = 0.286$$

Notes

Example 5.15: see Ex. 5.13 (2)

According to *Bayes' theorem* we obtain the other searched-for probabilities:

$$P(A_3|F) = \frac{0.3 \cdot 0.03}{0.035} = 0.257$$
$$P(A_4|F) = \frac{0.4 \cdot 0.02}{0.035} = 0.229$$

Calculations in R:

```
P_Ai_F <- P_Ai * P_F_Ai / P_F
round(P_Ai_F, 3)
[1] 0.229 0.286 0.257 0.229
```


Notes

Example 5.16: Missing probabilities (1)

A_1, A_2 and A_3 are three disjoint events that unite to Ω. Let $P(A_1) = 0.3$, $P(A_2) = 0.5$, $P(B|A_1) = 0.6$, $P(B|A_2) = 0.5$ and $P(B|A_3) = 0.1$.

Data input in R:

```
P_Ai   <- c(0.3,0.5,NA)
P_B_Ai <- c(0.6,0.5,0.1)
```

We calculate $P(A_1|B)$.

First we have $P(A_3) = 1 - 0.3 - 0.5 = 0.2$.

Calculations in R:

```
P_Ai[3] <- 1 - P_Ai[1] - P_Ai[2]
P_Ai
[1] 0.3 0.5 0.2
```


Notes

Example 5.16: Missing probabilities (2)

$$P(A_1|B) = \frac{0.3 \cdot 0.6}{0.3 \cdot 0.6 + 0.5 \cdot 0.5 + 0.2 \cdot 0.1} = \frac{0.18}{0.45} = \frac{2}{5} = 0.4$$

Calculations in R:

```
P_A1_B <- P_Ai[1] * P_B_Ai[1] / sum(P_Ai * P_B_Ai)
P_A1_B
[1] 0.4
```


Notes

Example 5.17: Accidents

We have the following notation:

W: Woman

M: Man

A: Accident

The probability for an accident is $P(A) = 1/100,000$. Furthermore $P(W|A) = 1/8$ and $P(M|A) = 7/8$.

Can we conclude from this that women are better drivers because they cause less accidents?

Notes

Urn model with replacement (WR)

Wirtschafts- und Sozialstatistik Universität Trier

In an urn there are N different balls.
From there we take n balls WR.

- Multiple draws possible
- **Independence** of draws

There are N^n variations.

Calculation of the number of variations in R:

```
# Attention: N and n must be defined
Number_of_variations <- N^n
```

Random draws following an urn model WR in R:

```
set.seed(123) # starting point for (pseudo-) random number
# generator
draws <- sample(x = 1:100, size = 3, replace = TRUE)
draws
[1] 31 79 51
```


Notes

Example 5.18: WR

Wirtschafts- und Sozialstatistik Universität Trier

In an urn there are 100 balls with the numbers 1 to 100. We draw three times one ball, note the number and return it into the urn.

$$P(K_1 = 31; K_2 = 79; K_3 = 51) = \frac{1}{100} \cdot \frac{1}{100} \cdot \frac{1}{100} = 10^{-6}$$

Notes

Urn model without replacement (WoR)

Wirtschafts- und Sozialstatistik Universität Trier

In an urn there are N different balls.
From this we draw n balls WoR.

- Multiple draws not possible (number of balls decreases)
- **Dependence** of the draws

There are $N \cdot (N-1) \cdot \dots \cdot (N-n+1) = \dfrac{N!}{(N-n)!}$ variations.

Calculation of the number of variations in R:

```
# Attention: N and n must be defined
Number_of_variantions <- factorial(N)/factorial(N-n)
```

Random draws following the urn model WoR in R:

```
set.seed(321) #  new starting point
draws <- sample(x = 1:100, size = 3, replace = FALSE)
draws
[1] 54 77 88
```


Notes

Example 5.19: WoR

Wirtschafts- und Sozialstatistik Universität Trier

In an urn there are 100 balls with the numbers 1 to 100. We draw three times one ball and note the number.

$$P(K_1 = 54; K_2 = 77; K_3 = 88) =$$
$$W(K_1 = 54) \cdot P(K_2 = 77 | K_1 = 54) \cdot P(K_3 = 88 | K_1 = 54, K_2 = 77) =$$
$$\frac{1}{100} \cdot \frac{1}{99} \cdot \frac{1}{98}$$

For $n/N < 0{,}05$ there is only a small numerical difference.

Notes

Wirtschafts- und Sozialstatistik
Universität Trier

More urn models

- WoR without ordering:
  As before, but here $n!$ orders of the balls are identical.

  There are $\dfrac{N!}{(N-n)! \cdot n!} = \binom{N}{n}$ combinations.

  Lottery as example: $\binom{49}{6} = 13,983,816$

  Determination in R:

```
Number_of_variantions <- choose(49,6)
```

- WR without ordering:

  There are $\binom{N+n-1}{n}$ combinations

  Determination in R:

```
# Attention: N and n must be defined
Number_of_variantions <- choose(N+n-1,n)
```


Notes

Notes

Notes

Notes