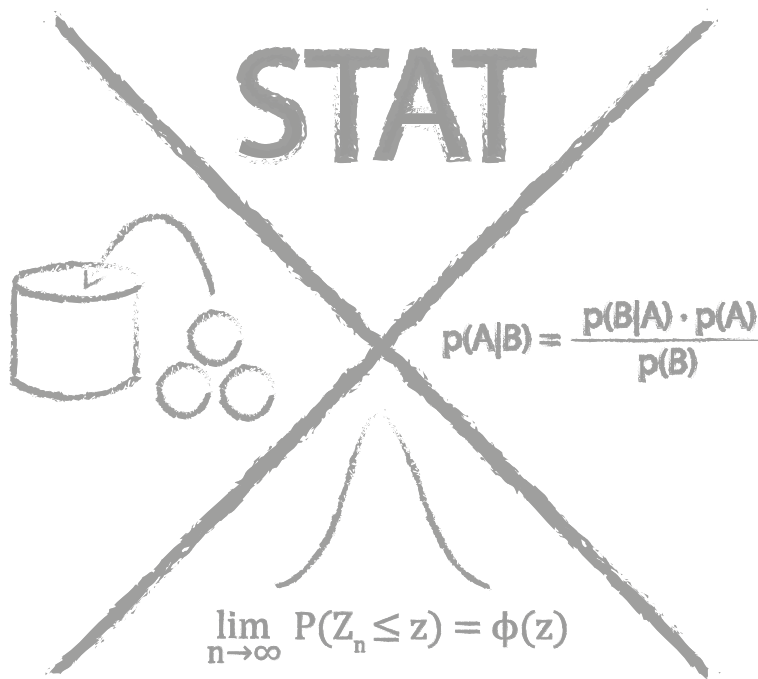


# Grundlagen Statistik

## - Übungsbuch -





Dieses Übungsbuch wird durch das Zentrum für Angewandte Ökonomik (kurz ZAÖ) der DHBW Ravensburg bereitgestellt.

Autoren: Prof. Dr. Daniel Blochinger  
Illustration: Prof. Dr. Daniel Blochinger  
Lizenz: [CC BY-NC-SA 4.0](#)

Weitere Lehr- und Lernmaterialien finden Sie auf unserer [Webseite](#).

Fehler gefunden? E-Mail an [blochinger@dhbw-ravensburg.de](mailto:blochinger@dhbw-ravensburg.de)!

## Aufgabe 1 Teilgebiete der Statistik

Zu welchem Teilgebiet der Statistik gehören folgende Aussagen?

- Der wahrscheinlichste Wurf bei Siedler von Catan (Summe zweier Würfel) ist die 7
- Von 1000 befragten Personen sind 32.1% CDU Wähler.
- Aus unserer Stichprobe schließen wir, dass Huskys signifikant größer sind als Chihuahwas.
- Der Notendurchschnitt in Statistik ist 2.4
- Die Wahrscheinlichkeit von Rot im Roulette ist 18/38

## Aufgabe 2 Grundbegriffe der deskriptiven Statistik

Überlege bei folgenden Datensätzen, ob die Merkmale quantitativ oder qualitativ sind und welche Eigenschaften sie haben (binär, diskret usw.).

Gebe bei quantitativen Merkmalen zusätzlich an, ob sie auf einer Ordinal- oder Kardinalskala gemessen werden.

Fahrzeug	NCAP	Norm	Vorbesitzer	TÜV?	Preis (€)
VW Golf 2.0 TSI	5	Euro 5	2	Ja	14990
VW Golf R Cabrio	3	Euro 4	1	Ja	22149
VW Golf GTI	5	Euro 5	1	Ja	18900
Mercedes E200	5	Euro 6a	3	Ja	24790
VW T6 Caravelle	5	Euro 6b	1	Ja	59980

Landkreis	Einwohner	Größe (m²)
Bodensee	218885	664
Biberach	203224	1409
Konstanz	288097	818
Ravensburg	287001	1632
Sigmaringen	131725	1204
Tuttlingen	142414	734

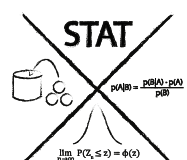
Land	BIP (Mrd.€)	Länder
Deutschland	3870	16
Österreich	447	9
Schweiz	677	26

## Aufgabe 3 Lageparameter

Berechne für die numerischen Merkmale der folgenden Datensätze jeweils Mittelwert, Median und Modalwert. Was fällt beim Vergleich der Lagemaße auf?

Name	Jahresbrutto
R. Müller	49500
H. Maier	74000
F. Probst	110500
C. Krumm	80500
J. Besos	89628000
S. Schiller	57500

Name	Punktzahl	Trefferquote
Andi Arbeit	7650	75%
Rainer Zufall	7230	70%
Klara Himmel	7920	80%
Andi Bar	6470	70%
Verleihnix	8430	90%
Kannix	460	5%



## Aufgabe 4 Lageparameter II

Bei einer Statistikklausur mit einem Maximum von 100 Punkten sind 7 von 8 Klausuren bereits korrigiert.

a) Berechnen Sie Mittelwert, Median und Modalwert für den aktuellen Korrekturstand.

b) Wie viel Punkte müsste der letzte Studierende erreichen, damit der Mittelwert 62.5 erreicht.

c) Welche Werte kann der Median noch erreichen?

d) Welche Werte kann der Modalwert noch erreichen?

Martikel Nr.	Klausurpkt.
564876	47
321890	72
121896	91
217866	18
125647	45
487963	67
654312	80
201405	???

## Aufgabe 5 Streuungsparameter

Unsere IT untersucht die Qualität der Internetverbindung an verschiedenen Standorten. Berechnen sie für das Merkmal Ping ...

a) ... Mittelwert und Median

b) ... Standardabweichung und lineare Abweichung

c) Berechne dieselben Kennzahlen ohne die Weinbergstraße. Was fällt bezüglich der Lage- und Streuparametern auf?

Messung	Ping (ms)
Marienplatz	15
Oberamtei	45
Marktstraße 28	30
Marktstraße 13	35
Rudolphstraße	25
Weinbergstraße	175
Klösterle	60

## Aufgabe 6 Streuungsparameter II

Berechne für die numerischen Merkmale der folgenden Datensätze jeweils Varianz, Standardabweichung und die Spannweite.

Athlet	Weite (m)	Punkte	Kennzeichen	Messung (km/h)	Bußgeld (€)
Kubacki	121.5	265.1	RV-XY-0815	66.4	97.50
Granerud	134.0	261.7	BC-00-666	89.3	288.50
Lanisek	121.5	258.8	RT-BD-1990	57.4	58.50
Kraft	125.0	255.0	ES-EL-1893	56.0	58.50
Stoch	121.0	249.2	GP-GP-2000	62.1	78.50
Lindvik	121.0	246.2	FN-HD-1980	64.8	78.50



## Aufgabe 7 Kovarianz und Korrelation

Berechne und interpretiere die Korrelation zwischen dem Eisumsatz und der Anzahl Sonnenstichen.

Visualisiere den Zusammenhang der beiden Größen in einem geeigneten Schaubild.

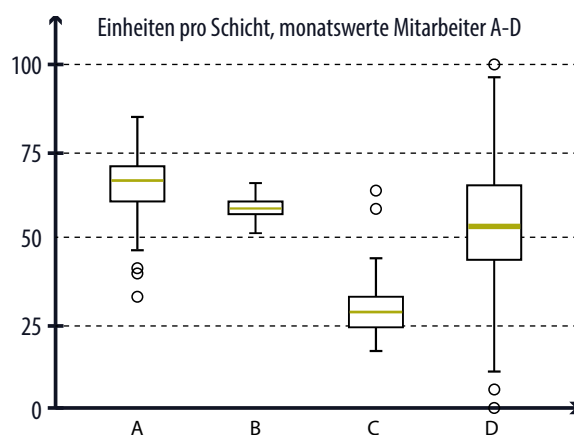
Temperatur (°C)	Umsatz Eis (€)	Sonnenstiche
18	900	8
24	1500	14
27	1800	25
32	1600	41
26	1700	32
11	0	0

## Aufgabe 8 Boxplot

Ein mittelständisches Unternehmen erfasst die Performance seiner Akkordarbeiter in der Fertigung.

Von jedem der vier Mitarbeiter wurden dabei an 30 Tage die Anzahl an korrekt gefertigten Einheiten gezählt.

Das Ergebnis wird Ihnen in Form eines Boxplots vorgelegt. Wie interpretieren Sie die Untersuchung?



## Aufgabe 9 Ernte Datensatz

Lade den auf Moodle bereitgestellten Datensatz „Ernte“ herunter. Verwende Excel oder falls du möchtest auch R bzw. SPSS, um folgende Aufgaben zu bearbeiten:

- Berechne für jedes Jahr die über alle Anbauprodukte summierte Anbauflächen und stelle deren absoluten Wert sowie die prozentuale Änderung zum Vorjahr in geeigneten Schaubildern dar.
- Zeige die Entwicklung von Produktion und Flächenertrag bei Kartoffeln, Karotten, Mais und Weizen. Erstelle für jedes Anbauprodukt ein Liniendiagramm, in dem die Entwicklung beider Merkmale gezeigt wird. Verwende ggf. eine zweite Achse, damit beide Linien gut ablesbar sind.
- Berechne den Korrelationskoeffizienten zwischen Produktionsfläche und Flächenertrag bei Kartoffeln, Karotten, Mais und Weizen.

## Aufgabe 10 YouTube Datensatz

Lade den auf Moodle bereitgestellten Datensatz „YouTube“ herunter. Dieser listet alle trendenden Videos aus dem Jahr 2018. Verwende Excel oder falls du möchtest auch R bzw. SPSS, um folgende Aufgaben zu bearbeiten:

- Berechne für das Merkmal Views Mittelwert, Median und Standardabweichung. Visualisiere die Verteilung mit einem Histogramm.
- Ergänze die Tabelle um eine Spalte, welche die Like/Dislike Ratio angibt. Berechne den Korrelationskoeffizienten zwischen der Anzahl der Kommentare und dieser Kennzahl.
- Zeige die relative Häufigkeit der Genres bzgl. trendenden Videos als Tabelle, als Balkendiagramm und als Tortendiagramm. Fasse bei den Diagrammen Genres mit weniger als 1% Anteil unter der Sammelkategorie „Sonstiges“ zusammen.

## Aufgabe 11 Elementarereignisse

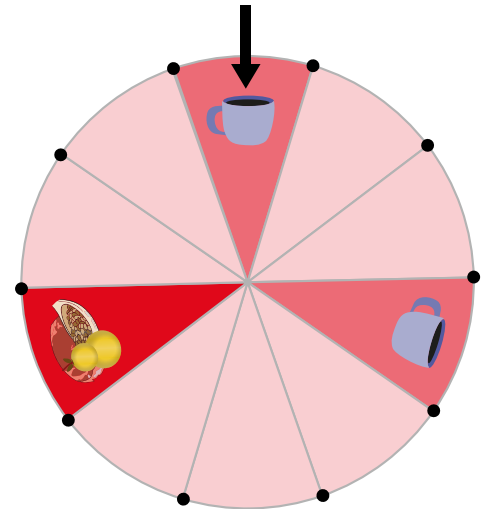
Gebe die Ereignismenge  $\Omega$  für folgende Zufallsereignisse an und kategorisieren diese in diskret und kontinuierlich. Gerne darfst du dabei Auslassungspunkte verwenden, also z. B. 1,2,3,...,100

- Du würfelst einen 20 seitigen Würfel.
- Du übersiehst eine Ampel. Ihre Farbe war ...
- Du pflanzt eine Paprikapflanze an der 3-8 Paprika wachsen
- Du wiegst die Paprikaernte, wobei jede Paprika zwischen 80 und 240 Gramm wiegen kann.

## Aufgabe 12      Wahrscheinlichkeiten & Ereignisse I

Das Studierendenwerk bietet allen Erstis ein Glücksradspiel in der Mensa an. Der Pfeil des Glücksrades zeigt am Ende auf eines von 10 Feldern, wobei der Ersti bei zwei Feldern einen kostenlosen Kaffee und bei einem Feld ein kostenloses Mittagessen gewinnt.

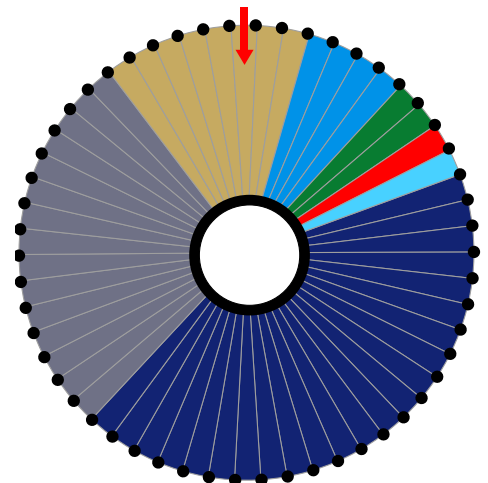
- Gib die Ereignismenge  $\Omega$  für einen einzelnen Spin an.
- Wie hoch ist die Wahrscheinlichkeit der Elementarereignisse?
- Wie hoch ist die Wahrscheinlichkeit, dass ein Ersti irgendetwas gewinnt? Definiere dazu auch ein entsprechendes Ereignis!
- Wie hoch ist die Wahrscheinlichkeit, dass das Studierendenwerk nach 10 Spins mindestens einen Gewinn ausgeben muss?



## Aufgabe 13      Wahrscheinlichkeiten & Ereignisse II

Wir betrachten ein klassisches Glücksradspiel, bei dem der Pfeil am Ende auf eines von 54 Felder zeigt. Die Felder sind mit sieben verschiedenen, unterschiedlich häufigen Edelsteinsymbolen versehen.

- Gib die Ereignismenge  $\Omega$  für einen einzelnen Spin an.
- Wie hoch ist die Wahrscheinlichkeit, dass der Pfeil auf einen Topas zeigt?
- Wie hoch ist die Wahrscheinlichkeit, dass der Pfeil auf einen Rubin oder einen Diamanten zeigt? Definiere dazu auch ein entsprechendes Ereignis!
- Wie hoch ist die Wahrscheinlichkeit, dass dieses Ereignis zweimal in Folge eintritt?
- Wie hoch ist die Wahrscheinlichkeit, dass dieses Ereignis zehnmal in Folge NICHT eintritt?
- Wie hoch ist die Wahrscheinlichkeit, dass bei 5 Spins mindestens ein Amethyst auftritt?



23x Saphir	2x Smaragd
15x Amethyst	1x Rubin
8x Topas	1x Diamant
4x Opal	$\Sigma = 54$ Felder

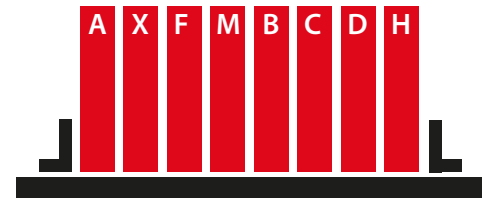
## Aufgabe 14 Permutationen und Binomialkoeffizienten

a) Sie bestücken ein frisch montiertes Regal mit acht Büchern. Wie viele mögliche Anordnungen gibt es, wenn Sie genau acht Bücher dabei zur Auswahl haben?

b) Wie viele Auswahlmöglichkeiten gibt es, wenn Sie zwölf Bücher zur Auswahl haben aus denen Sie acht für das Regal auswählen? Wie viele verschiedene Anordnungen an Büchern gibt es insgesamt?

c) Ein Bogenschütze trifft ein Ziel mit einer Wahrscheinlichkeit von 80%. Berechnen Sie die Wahrscheinlichkeiten, dass er von 3 Schüssen genau 0,1,2 bzw. 3 Schüsse trifft.

d) In einer Packung „Das gute Himbeermüsli“ vom Discounter Brutto sollte genau eine Himbeere enthalten sein. Durch einen Produktionsfehler fehlt diese Himbeere jedoch in jeder 100-sten Packung. Die Verbraucherschutzzentrale führt eine Stichprobe mit 20 Packungen durch. Findet Sie mindestens zwei Packungen ohne Himbeere, kann Sie den Discounter belangen. Wie hoch ist die Wahrscheinlichkeit, dass ihr dies gelingt?



## Aufgabe 15 Urnenmodelle I

Wir betrachten das Glücksspiel KENO-10. Auf dem KENO-Schein kreuzen wir 10 von 70 Zahlen an. Gezogen werden insgesamt 20 Gewinnzahlen.

a) Wie hoch ist die Wahrscheinlichkeit, dass alle 10 getippten Zahlen eine Gewinnzahl sind?

b) Wie hoch ist die Wahrscheinlichkeit, dass genau 9 von unseren 10 getippten Zahlen Gewinnzahlen sind?

c) Wie hoch ist die Wahrscheinlichkeit, dass keine einzige der 10 getippten Zahlen eine Gewinnzahl sind?



## Aufgabe 16 Urnenmodelle II

Betrachten Sie ein elektronisches Zahlenschloss, bei dem wir mit den Ziffern 0-9 einen fünfstelligen Code eingeben müssen.



- Wie hoch ist die Wahrscheinlichkeit, dass wir mit einem Versuch genau den richtigen Code eingeben?
- Wie viele verschiedene Kombinationen müssen wir ausprobieren, damit wir zu 50% Wahrscheinlichkeit den richtigen Code finden?
- Wie viele sind es, wenn wir uns nicht merken, welche Kombinationen wir bereits erfolglos probiert haben?
- Der Programmierer des Schlosses hat einen Fehler gemacht. Es öffnet die Türe, wenn von den 5 Ziffern mindestens 4 richtig sind. Wie hoch ist die Wahrscheinlichkeit die Tür im ersten Versuch aufzubekommen?
- Der Kundendienst behebt den Fehler der letzten Teilaufgabe, baut dabei aber versehentlich einen neuen Fehler ein: Jetzt spielt die Reihenfolge der Ziffern keine Rolle mehr, d.h. wenn 10125 der richtige Code wäre, würden Sie auch mit 25110 die Tür öffnen können. Wie hoch ist die Wahrscheinlichkeit die Tür im ersten Versuch aufzubekommen?
- Nachdem das Schloss endlich fehlerfrei programmiert ist, tut sich ein anderes Problem auf: Durch die Benutzung der Eingabetasten sind die Zahlen aus denen der Code besteht offensichtlich: 0,4,7,8 und 9. Wie hoch ist die Wahrscheinlichkeit die Tür im ersten Versuch aufzubekommen?

## Aufgabe 17 Urnenmodelle III

Bei der Pokervariante Five-Card-Draw erhält jeder Spieler fünf Handkarten aus einem 52-er Kartendeck.

- Wie hoch ist die Wahrscheinlichkeit eines Flushes, d. h. alle fünf Karten mit demselben Symbol?
- Wie hoch ist die Wahrscheinlichkeit eines Vierlings, d. h. vier Karten mit demselben Wert?
- Wie hoch ist die Wahrscheinlichkeit eines Full House, d. h. drei Karten mit demselben Wert und zwei Karten mit einem anderen gleichen Wert?

♥ 2	♣ 2	♦ 2	♠ 2
♥ 3	♣ 3	♦ 3	♠ 3
♥ 4	♣ 4	♦ 4	♠ 4
♥ 5	♣ 5	♦ 5	♠ 5
♥ 6	♣ 6	♦ 6	♠ 6
♥ 7	♣ 7	♦ 7	♠ 7
♥ 8	♣ 8	♦ 8	♠ 8
♥ 9	♣ 9	♦ 9	♠ 9
♥ 10	♣ 10	♦ 10	♠ 10
♥ J	♣ J	♦ J	♠ J
♥ Q	♣ Q	♦ Q	♠ Q
♥ K	♣ K	♦ K	♠ K
♥ A	♣ A	♦ A	♠ A

## Aufgabe 18 Urnenmodelle IV

Betrachten Sie eine Bar mit 20 verschiedenen Spirituosen, Säften und Sirupen, darunter auch die rechts auf dem Rezept genannten.

a) Sie nehmen einen 20ml Messbecher und füllen ihn 8 Mal mit einer zufällig ausgewählten Flüssigkeit. Wie hoch ist die Wahrscheinlichkeit, dass sie zufällig einen Long Island Iced Tea dabei mischen?

### Long Island Iced Tea Rezept

20ml Tequilla  
20ml Vodka  
20ml White Rum  
20ml Cointreau  
20ml Gin  
40ml Zitronensaft  
20ml Zuckersirup

## Aufgabe 19 Bedingte Wahrscheinlichkeit I

Ein Angler geht jeden Sonntagabend zum Angeln. Bei guten Wetterbedingungen fängt er zu 75% einen Fisch, bei schlechten Wetterbedingungen nur zu 40%. Die Wahrscheinlichkeit für ein geeignetes Wetter liegt bei 60%.

- Definiere die Zufallsereignisse und schreibe die gegebenen Wahrscheinlichkeiten als Gleichungen.
- Berechne die Wahrscheinlichkeit, dass der Angler am nächsten Sonntag einen Fisch fängt.
- Stelle alle Ereignisse und Wahrscheinlichkeiten in Form eines Wahrscheinlichkeitsbaums dar.

## Aufgabe 20 Bedingte Wahrscheinlichkeit II

Ein Agrarhändler verkauft ein „Trächtigkeitsprüfgerät“ für Rinder, das ein trächtiges Rind mit 95% Sicherheit feststellen kann und ein nicht-trächtiges Rind zu 99% als nicht trächtig einstuft. Der Test soll verhindern, dass trächtige Rinder geschlachtet werden, denn laut Statistik sind 5% aller geschlachteter Rinder trächtig gewesen.

- Definiere die Zufallsereignisse und schreibe die gegebenen Wahrscheinlichkeiten als Gleichungen.
- Wie hoch ist die Wahrscheinlichkeit, dass ein als trächtig angezeigtes Rind auch trächtig ist?
- Wie hoch ist die Wahrscheinlichkeit trotz negativen Tests ein trächtiges Rind zu schlachten?

### Aufgabe 21 Bedingte Wahrscheinlichkeit III

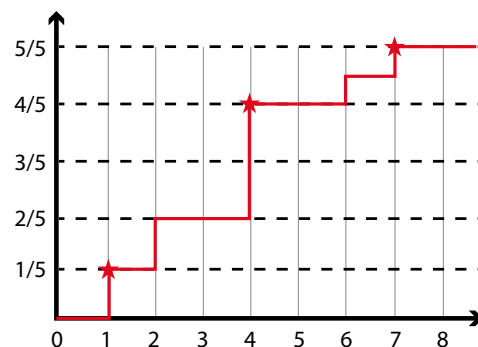
„6 von 10 tödlich verunglückten Autofahrern trugen einen Sicherheitsgurt“ - Um diese Headline besser einschätzen zu können recherchieren sie, dass 99% aller Autofahrer einen Sicherheitsgurt tragen. Eine Wahrscheinlichkeit für einen tödlichen Unfall können Sie jedoch nicht finden.

- Definiere die beiden Zufallsereignisse und stelle die gegebenen Wahrscheinlichkeiten formal dar.
- Versuche mit den gegebenen Wahrscheinlichkeiten herauszufinden, ob Sicherheitsgurte positiv, neutral oder negativ für die Sicherheit des Fahrers sind.

### Aufgabe 22 Zufallsvariablen I

Betrachten Sie die rechts gezeigte folgende Verteilungsfunktion der Zufallsvariable X!

- Geben Sie die Bedeutung der drei markierten Punkte verbal an.
- Wie hoch ist die Wahrscheinlichkeit, dass X mindestens 2.5 ist?
- Zeichne eine zur Verteilungsfunktion passende Dichtefunktion!
- Berechne den Erwartungswert von X.



### Aufgabe 23 Zufallsvariablen II

Wir betrachten einen vierseitigen Würfel, der mit einer Wahrscheinlichkeit von jeweils 1/4 die Augenzahlen 0, 1, 2 und 3 zeigt. Die Zufallsvariable X entspricht der gewürfelten Augenzahl.

- Stelle die Verteilung als Verteilungsfunktion und als Wahrscheinlichkeitsfunktion dar.
- Berechne den Erwartungswert und die Varianz von X.

### Aufgabe 24 Zufallsvariablen III

Zwei Zufallsgeneratoren erzeugen gleichverteilte Zahlen auf den Intervallen -1 bis 1 und 10 bis 20. Die Zufallsvariable X (-1 bis 1) und Y (10 bis 20) sind die jeweils generierten Zahlen.

- Stelle die Verteilung von X und Y als Verteilungsfunktion und als Dichtefunktion dar.
- Berechne den Erwartungswert von X, Y und X+Y sowie die Varianz von X und Y.

## Aufgabe 25 Zufallsvariablen IV

Wir betrachten zwei verschiedene Anbieter von Rubbellosen: KingCash und BigWin. Die Tabelle rechts gibt die Häufigkeit von Nieten und Gewinnlosen an. Beide Anbieter verlangen 1€ pro Los.

Die Zufallsvariable X beschreibt die Auszahlung eines KingCash Loses, die Zufallsvariable Y beschreibt dagegen die Auszahlung eines BigWin Loses.

Von 1000 Losen sind...			
KingCash		BigWin	
800x	Nieten	900x	Nieten
100x	2€	50x	2€
50x	3€	25x	6€
50x	5€	25x	10€

- Stelle die Verteilung von X und Y als Verteilungsfunktion und als Wahrscheinlichkeitsfunktion dar.
- Berechne den Erwartungswert und die Varianz von X und Y
- Berechne den Erwartungswert und die Varianz von  $5X+5Y$ , d. h. die Auszahlung aus dem Kauf von jeweils fünf KingCash Losen und fünf BigWin Losen.

## Aufgabe 26 Z-Test

Die beiden Anbieter KingCash und BigWin werden einer Untersuchung unterzogen, da Zweifel bestehen, ob die Anzahl an Nieten nicht deutlich höher ist als in der Tabelle rechts oben angegeben.

Die zuständige Behörde kauft von jedem Anbieter 500 Lose und findet 415 Nieten bei KingCash und 470 Nieten bei BigWin.

Verwende einen einseitigen Z-Test, um einschätzen zu können, ob sich die beiden Hersteller auf juristische Konsequenzen einstellen sollten. Stelle dazu geeignete Hypothesenpaare auf und prüfe, ob die zu große Anzahl an Nieten in der Stichprobe signifikant sind oder als Zufall eingeordnet werden könnten.

## Aufgabe 27 T-Test

Lade den auf Moodle bereitgestellten Datensatz „Saatgut“ herunter. Verwende Excel oder falls du möchtest auch R bzw. SPSS, um die dort enthaltenen Aufgabenstellungen zu lösen.



## Aufgabe 1 Teilgebiete der Statistik

Wir teilen die Statistik in **d**eskriptive Statistik, **i**nduktive Statistik und **K**ombinatorik ein.

- K** Der wahrscheinlichste Wurf bei Siedler von Catan (Summe zweier Würfel) ist die 7
- D** Von 1000 befragten Personen sind 32.1% CDU Wähler
- I** Aus unserer Stichprobe schließen wir, dass Huskys signifikant größer sind als Chiawawas.
- D** Der Notendurchschnitt in Statistik ist 2.4
- K** Die Wahrscheinlichkeit von Rot im Roulette ist 18/38

## Aufgabe 2 Grundbegriffe der deskriptiven Statistik

Die Musterlösung wird in der Tabelle rechts gezeigt.

Die Eigenschaft der Skala wird nur bei quantitativen Merkmalen angegeben.

„Ordinal“ kann sowohl die Eigenschaft eines qualitativen Merkmals als auch die Art einer Skala sein.

Bei Preisen & BIP könnte man sich streiten, ob das Merkmal nicht doch diskret ist mit einem Cent als kleinste Einheit.

Merkmal	Quantitativ	Eigenschaft	Skala
Fahrzeug	✗	kategorial	-
NCAP	✓	diskret	ordinal
Abgasnorm	✗	ordinal	-
Vorbesitzer	✓	diskret	kardinal
TÜV	✗	binär	-
Preis	✓	kontinuierlich	kardinal
Einwohner	✓	diskret	kardinal
Fläche	✓	kontinuierlich	kardinal
BIP	✓	kontinuierlich	kardinal
Länder	✓	diskret	kardinal

## Aufgabe 3 Lageparameter

Berechne für die numerischen Merkmale der folgenden Datensätze jeweils Mittelwert, Median und Modalwert. Was fällt beim Vergleich der Lagemaße auf?

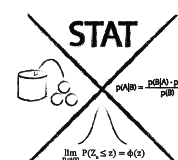
Name	Jahresbrutto
R. Müller	49500
H. Maier	74000
F. Probst	110500
C. Krumm	80500
J. Besos	89628000
S. Schiller	57500

└─→ a

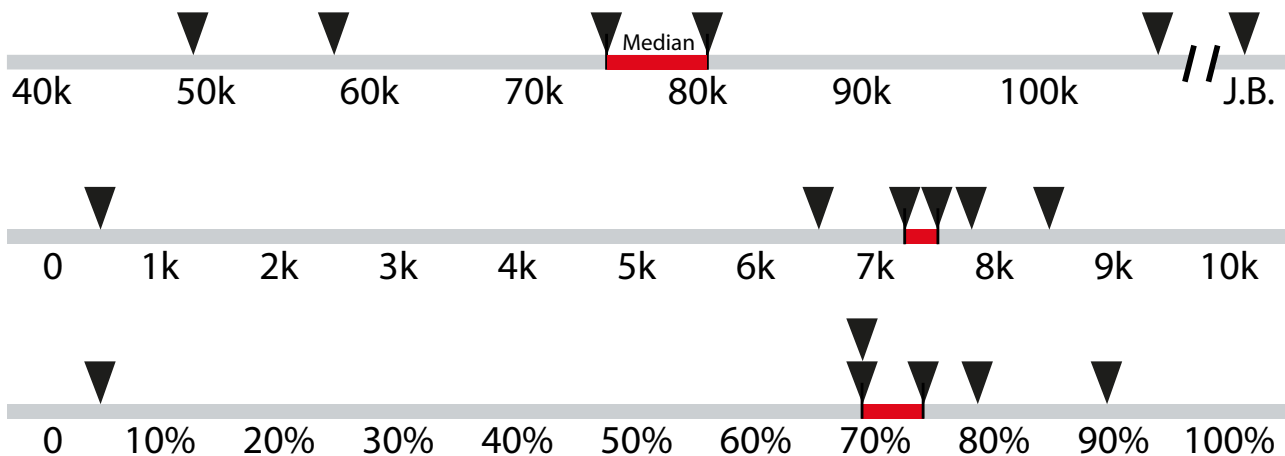
Name	Punktzahl	Trefferquote
Andi Arbeit	7650	75%
Rainer Zufall	7230	70%
Klara Himmel	7920	80%
Andi Bar	6470	70%
Verleihnix	8430	90%
Kannix	460	5%

└─→ b

└─→ c



$$\begin{aligned}\bar{a} &= \frac{1}{6} \sum_{i=1}^6 a_i & \tilde{a} &\in (74000, 80500) \\ &= \frac{1}{6} (49500 + 74000 + 110500 + 80500 + 89628000 + 57500) & a_m &\text{existiert nicht} \\ &= 15000000 \\ \bar{b} &= \frac{1}{6} \sum_{i=1}^6 b_i & \tilde{b} &\in (7230, 7650) \\ &= \frac{1}{6} (7650 + 7230 + 7920 + 6470 + 8430 + 460) & b_m &\text{existiert nicht} \\ &= 6360 \\ \bar{c} &= \frac{1}{6} \sum_{i=1}^6 c_i & \tilde{c} &\in (0.70, 0.75) \\ &= \frac{1}{6} (0.75 + 0.70 + 0.80 + 0.70 + 0.90 + 0.05) & c_m &= 0.70 \\ &= 0.65 = 65\%\end{aligned}$$



Da wir mit  $n=6$  eine gerade Anzahl an Beobachtungen haben, kann jeder Wert zwischen dem drittgrößten und drittkleinsten Wert als Median angegeben werden.

Es fällt auf, dass der Median robuster gegen Ausreißer ist als der Mittelwert. Beim ersten Merkmal ist der Mittelwert z. B. völlig nutzlos.

Auch der Modalwert lässt uns zweimal im Stich, wobei das an den kleinen Datensätzen liegt.

# Aufgabe 4      Lageparameter II

Bei einer Statistikklausur mit einem Maximum von 100 Punkten sind 7 von 8 Klausuren bereits korrigiert.

a) Berechnen Sie Mittelwert, Median und Modalwert für den aktuellen Korrekturstand.

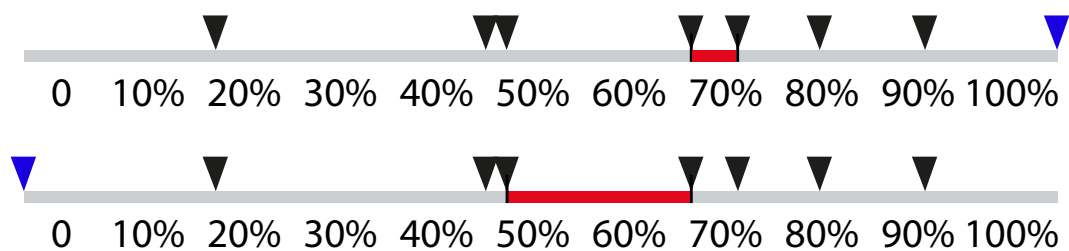
$$\begin{aligned}\bar{x} &= \frac{1}{7} \sum_{i=1}^7 x_i \\ &= \frac{1}{7} (47 + 72 + 91 + 18 + 45 + 67 + 80) \\ &= \frac{420}{7} = 60.0 \\ \tilde{x} &= 67\end{aligned}$$

Martikel Nr.	Klausurpkt.
564876	47
321890	72
121896	91
217866	18
125647	45
487963	67
654312	80
201405	???

b) Wie viel Punkte müsste der letzte Studierende erreichen, damit der Mittelwert 62.5 erreicht.

$$\begin{aligned}\bar{x} &= \frac{420 + x_8}{8} \stackrel{!}{=} 62.5 \quad | \cdot 8 \\ \Leftrightarrow 420 + x_8 &= 500 \quad | -420 \\ \Leftrightarrow x_8 &= 80\end{aligned}$$

c) Welche Werte kann der Median noch erreichen? Siehe markierte Flächen:



d) Welche Werte kann der Modalwert noch erreichen?

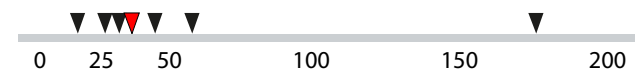
Der Modalwert kann nur noch die Werte erreichen, die bereits in der Tabelle vorhanden sind. Ergibt sich bei der Korrektur eine Punktzahl, die noch nicht in der Tabelle vorkommt, gibt es keinen Modalwert.

## Aufgabe 5      Streuungsparameter

Unsere IT untersucht die Qualität der Internetverbindung an verschiedenen Standorten. Berechnen sie für das Merkmal Ping ...

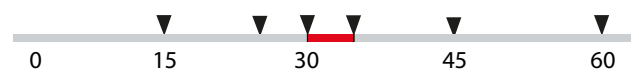
- a) ... Mittelwert und Median  
 b) ... Standardabweichung und lineare Abweichung  
 c) Berechne dieselben Kennzahlen ohne die Weinbergstraße. Was fällt bezüglich der Lage- und Streuparametern auf?

Messung	Ping (ms)
Marienplatz	15
Oberamtei	45
Marktstraße 28	30
Marktstraße 13	35
Rudolphstraße	25
Weinbergstraße	175
Klösterle	60



$$\begin{aligned}\bar{x} &= \frac{1}{7} \sum_{i=1}^7 x_i \\ &= \frac{1}{7} (15+45+30+35+25+175+60) \\ &= 55 \\ \tilde{x} &= 35\end{aligned}$$

$$\begin{aligned}\sigma^2 &= \frac{1}{6} \sum_{i=1}^7 (x_i - \bar{x})^2 \\ &= \frac{1}{6} [(15-55)^2 + (45-55)^2 + (30-55)^2 \\ &\quad (35-55)^2 + (25-55)^2 + (175-55)^2 \\ &\quad (60-55)^2] \\ &= \frac{1}{6} [(-40)^2 + (-10)^2 + (-25)^2 + (-20)^2 \\ &\quad (-30)^2 + (120)^2 + (5)^2] \\ &= \frac{1}{6} [1600+100+625+400+900 \\ &\quad + 14400 + 25] = 3008 \\ \sigma &= \sqrt{3008} = 54.85\end{aligned}$$



$$\begin{aligned}\bar{x} &= \frac{1}{6} \sum_{i=1}^6 x_i \\ &= \frac{1}{6} (15+45+30+35+25+60) \\ &= 35 \\ \tilde{x} &= 32.5 \text{ (Wert zwischen 30 und 35)}\end{aligned}$$

$$\begin{aligned}\sigma^2 &= \frac{1}{5} \sum_{i=1}^6 (x_i - \bar{x})^2 \\ &= \frac{1}{5} [(15-35)^2 + (45-35)^2 + (30-35)^2 \\ &\quad (35-35)^2 + (35-55)^2 + (60-35)^2] \\ &= \frac{1}{5} [(-20)^2 + (10)^2 + (-5)^2 + (-0)^2 \\ &\quad (-20)^2 + (25)^2] \\ &= \frac{1}{5} [400 + 100 + 25 + 400 + 625] \\ &= 250 \\ \sigma &= \sqrt{250} = 15.8\end{aligned}$$

$$\begin{aligned}
 MD &= \frac{1}{7} \sum_{i=1}^7 |x_i - \bar{x}| \\
 &= \frac{1}{7} [ |15-55| + |45-55| + |30-55| \\
 &\quad |35-55| + |25-55| + |175-55| \\
 &\quad |60-55| ] \\
 &= \frac{1}{7} [ 40+10+25+20+30+120+5 ] \\
 &= 35.7
 \end{aligned}$$

$$\begin{aligned}
 MD &= \frac{1}{6} \sum_{i=1}^6 |x_i - \bar{x}| \\
 &= \frac{1}{6} [ |15-35| + |45-35| + |30-35| \\
 &\quad |35-35| + |25-35| + |60-35| ] \\
 &= \frac{1}{6} [ 20+10+5+0+10+25 ] \\
 &= 11.67
 \end{aligned}$$

Bei den Lageparametern ist der Median robuster gegen Ausreißer als der Mittelwert. Bei den Streuungsparametern ist die lineare Abweichung robuster als die Standardabweichung.

### Aufgabe 6      Streuungsparameter II

Berechne für die numerischen Merkmale der folgenden Datensätze jeweils Varianz, Standardabweichung und die Spannweite.

Athlet	Weite (m)	Punkte	Kennzeichen	Messung (km/h)	Bußgeld (€)
Kubacki	121.5	265.1	RV-XY-0815	66.4	97.50
Granerud	134.0	261.7	BC-00-666	89.3	288.50
Lanisek	121.5	258.8	RT-BD-1990	57.4	58.50
Kraft	125.0	255.0	ES-EL-1893	56.0	58.50
Stoch	121.0	249.2	GP-GP-2000	62.1	78.50
Lindvik	121.0	246.2	FN-HD-1980	64.8	78.50

$\xrightarrow{\quad} w \quad \xrightarrow{\quad} s \quad \xrightarrow{\quad} v \quad \xrightarrow{\quad} p$

Die Spannweiten können wir direkt bestimmen:

$$R_w = \max(w_i) - \min(w_i) = 134.0 - 121.0 = 13.0$$

$$R_s = \max(w_s) - \min(w_s) = 265.1 - 246.2 = 18.9$$

$$R_v = \max(w_v) - \min(w_v) = 89.3 - 56.0 = 33.3$$

$$R_p = \max(w_p) - \min(w_p) = 288.50 - 58.50 = 230$$

Für die anderen Streuungsparameter benötigen wir die Mittelwerte der Merkmale:

$$\begin{aligned}\bar{w} &= \frac{1}{6} \sum_{i=1}^6 w_i \\ &= \frac{1}{6} (121.5 + 134.0 + 121.5 + 125.0 + 121.0 + 121.0) \\ &= 124.0\end{aligned}$$

$$\begin{aligned}\bar{s} &= \frac{1}{6} \sum_{i=1}^6 s_i \\ &= \frac{1}{6} (265.1 + 261.7 + 258.8 + 255.0 + 249.2 + 246.2) \\ &= 256.0\end{aligned}$$

$$\begin{aligned}\bar{v} &= \frac{1}{6} \sum_{i=1}^6 v_i \\ &= \frac{1}{6} (66.4 + 89.3 + 57.4 + 56.0 + 62.1 + 64.8) \\ &= 66.0\end{aligned}$$

$$\begin{aligned}\bar{p} &= \frac{1}{6} \sum_{i=1}^6 p_i \\ &= \frac{1}{6} (97.5 + 288.5 + 58.5 + 58.5 + 78.5 + 78.5) \\ &= 110.0\end{aligned}$$

Wir berechnen die Varianzen und daraus auch die Standardabweichung:

$$\begin{aligned}\sigma_w^2 &= \frac{1}{5} \sum_{i=1}^6 (w_i - \bar{w})^2 & \sigma_w &= \sqrt{26.3} = 5.13 \\ &= \frac{1}{5} [ (121.5 - 124.0)^2 + (133.0 - 124.0)^2 + \dots ] \\ &= 26.3\end{aligned}$$

$$\begin{aligned}\sigma_s^2 &= \frac{1}{5} \sum_{i=1}^6 (s_i - \bar{s})^2 & \sigma_s &= \sqrt{53.28} = 7.30 \\ &= \frac{1}{5} [ (265.2-256.0)^2 + (261.7-256.0)^2 + \dots ] \\ &= 53.28\end{aligned}$$

$$\begin{aligned}\sigma_v^2 &= \frac{1}{5} \sum_{i=1}^6 (v_i - \bar{v})^2 & \sigma_v &= \sqrt{146.73} = 12.11 \\ &= \frac{1}{5} [ (67.0-66.0)^2 + (89.3-66.0)^2 + \dots ] \\ &= 146.73\end{aligned}$$

$$\begin{aligned}\sigma_p^2 &= \frac{1}{5} \sum_{i=1}^6 (p_i - \bar{p})^2 & \sigma_p &= \sqrt{7861.5} = 88.67 \\ &= \frac{1}{5} [ (98.5-110.2)^2 + (288.5-110.2)^2 + \dots ] \\ &= 7861.5\end{aligned}$$

### Aufgabe 7 Kovarianz und Korrelation

Berechne und interpretiere die Korrelation zwischen dem Eisumsatz und der Anzahl Sonnenstichen.

Visualisiere den Zusammenhang der beiden Größen in einem geeigneten Schaubild.

Temperatur (°C)	Umsatz Eis (€)	Sonnenstiche
18	900	8
24	1500	14
27	1800	25
32	1600	41
26	1700	32
11	0	0

$$\begin{aligned}\overline{\text{eis}} &= \frac{1}{6} \sum_{i=1}^6 \text{eis}_i \\ &= \frac{1}{6} (900 + 1500 + 1800 + 1600 + 1700 + 0) \\ &= 1250\end{aligned}$$

$$\begin{aligned}\overline{\text{sst}} &= \frac{1}{6} \sum_{i=1}^6 \text{sst}_i \\ &= \frac{1}{6} (8 + 14 + 25 + 41 + 32 + 0) \\ &= 20\end{aligned}$$

$$\begin{aligned}\sigma_{\text{eis}}^2 &= \frac{1}{5} \sum_{i=1}^6 (\text{eis}_i - \overline{\text{eis}})^2 & \sigma_{\text{eis}} &= \sqrt{475000} = 689.2 \\ &= \frac{1}{5} [ (900-1250)^2 + (1500-1250)^2 + (1800-1250)^2 + \dots ] \\ &= 475000\end{aligned}$$

$$\begin{aligned}\sigma_{\text{sst}}^2 &= \frac{1}{5} \sum_{i=1}^6 (\text{sst}_i - \overline{\text{sst}})^2 & \sigma_{\text{sst}} &= \sqrt{238} = 15.4 \\ &= \frac{1}{5} [ (8-20)^2 + (14-20)^2 + (25-20)^2 + (41-20)^2 + \dots ] \\ &= 238\end{aligned}$$

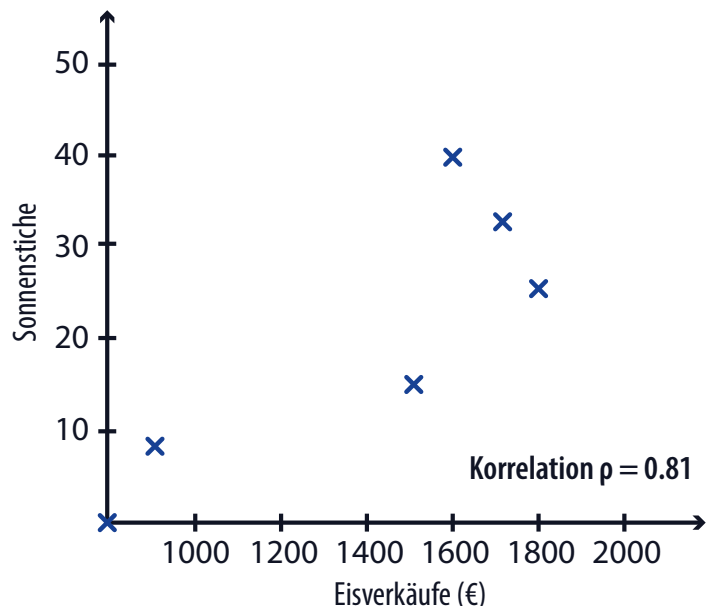
$$\begin{aligned}q_{\text{eis,sst}} &= \frac{1}{5} \sum_{i=1}^6 (\text{eis}_i - \overline{\text{eis}}) \cdot (\text{sst}_i - \overline{\text{sst}}) \\ &= \frac{1}{5} [ (900-1250)(8-20) + (1500-1250)(14-20) + \dots ] \\ &= 8640\end{aligned}$$

$$\rho_{\text{eis,sst}} = \frac{8640}{689.2 \cdot 15.4} = 0.81$$

Obwohl die Korrelation mit 0.81 sehr stark ist, kann man sie nicht kausal deuten.

Sehr wahrscheinlich ist hier eine dritte Einflussgröße im Spiel, welche Auswirkungen auf die beiden Merkmale „Eisumsatz“ und „Sonnenstiche“ hat. Beispiele: UV-Index, Sonnenscheindauer, Temperatur usw.

Wichtig: Die Entscheidung ob es zur Korrelation eine plausible Kausalität gibt oder nicht ist nicht von den Daten abhängig.





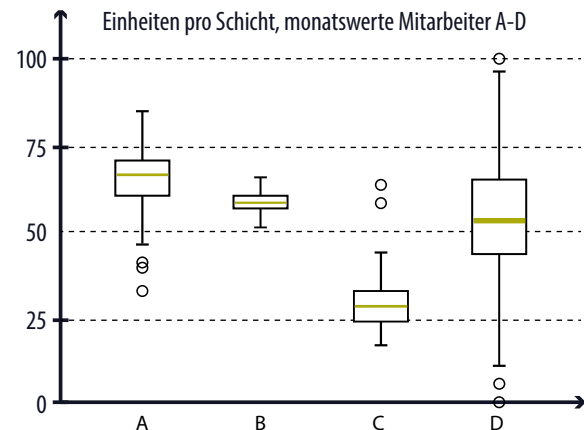
## Aufgabe 8 Boxplot

**Mitarbeiter A** hat den höchsten Median, aber auch eine hohe Varianz und einige besonders schlechte Tage.

**Mitarbeiter B** ist im Schnitt etwas langsamer als Mitarbeiter A, dafür jedoch sehr zuverlässig.

**Mitarbeiter C** scheint grundsätzlich langsam zu sein, hat jedoch zwei Ausreiser nach oben, bei denen er mit den anderen mithalten kann. Mögliche Hypothese: Er arbeitet langsamer, als er könnte und fühlte sich an den beiden Ausreiser Tagen beobachtet.

**Mitarbeiter D** ist durch seine hohe Varianz völlig unzuverlässig und das, obwohl er den Rekord von 100 Teilen in einer Schicht hält.



## Aufgabe 11 Elementarereignisse

Gebe die Ereignismenge  $\Omega$  für folgende Zufallsereignisse an und kategorisieren diese in diskret und kontinuierlich. Gerne darfst du dabei Auslassungspunkte verwenden, also z. B. 1,2,3,...,100

a) Du würfelst einen 20 seitigen Würfel.

Diskret mit  $\Omega = \{1,2,3,\dots,20\}$

b) Du übersiehst eine Ampel. Ihre Farbe war ...

Diskret mit  $\Omega = \{\text{rot, gelb, grün}\}$

c) Du pflanzt eine Paprikapflanze an der 3-8 Paprika wachsen

Diskret mit  $\Omega = \{3,4,5,6,7,8\}$

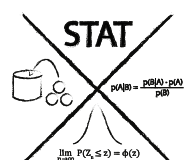
d) Du wiegst die Paprikaernte, wobei jede Paprika zwischen 80 und 240 Gramm wiegen kann.

Kontinuierlich mit  $\Omega = [80,240]$  für einzelne Paprika

Kontinuierlich mit  $\Omega = [240,1920]$  für ganze Ernte

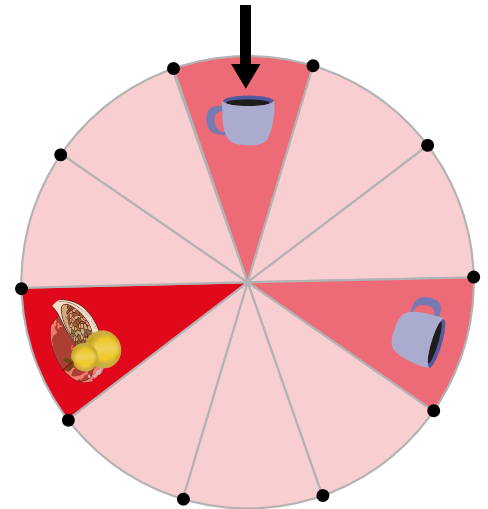


Wo sind die Lösungen für die Aufgaben 9 bis 11 geblieben? Bei diesen Aufgaben sind die Lösungen direkt in das zugehörige Excel-Sheet integriert!



## Aufgabe 12      Wahrscheinlichkeiten &amp; Ereignisse I

Das Studierendenwerk bietet allen Erstis ein Glücksradspiel in der Mensa an. Der Pfeil des Glücksrades zeigt am Ende auf eines von 10 Feldern, wobei der Ersti bei zwei Feldern einen kostenlosen Kaffee und bei einem Feld ein kostenloses Mittagessen gewinnt.



a) Gib die Ereignismenge  $\Omega$  für einen einzelnen Spin an.

Diskret mit  $\Omega = \{\text{Niete}, \text{Kaffee}, \text{Essen}\}$

b) Wie hoch ist die Wahrscheinlichkeit der Elementarereignisse?

$$P(\text{Niete}) = \frac{7}{10} = 0.7 = 70\%$$

$$P(\text{Kaffee}) = \frac{2}{10} = 0.2 = 20\%$$

$$P(\text{Essen}) = \frac{1}{10} = 0.1 = 10\%$$

c) Wie hoch ist die Wahrscheinlichkeit, dass ein Ersti irgendetwas gewinnt? Definiere dazu auch ein entsprechendes Ereignis!

Ereignis  $A = \{\text{Kaffee}, \text{Essen}\}$

$$P(A) = P(\text{Kaffee}) + P(\text{Essen}) = \frac{2}{10} + \frac{1}{10} = 30\%$$

d) Wie hoch ist die Wahrscheinlichkeit, dass das Studierendenwerk nach 10 Spins mindestens einen Gewinn ausgeben muss?

Ereignis  $A_i = \text{Ein Gewinn (Kaffee oder Essen) im Spin Nr. } i$

$$P(A_1 \cup \dots \cup A_{10}) = 1 - P(\overline{A}_1) \cdot \dots \cdot P(\overline{A}_{10}) = 1 - 0.7^{10} = 97.2\%$$

### Aufgabe 13      Wahrscheinlichkeiten & Ereignisse II

Wir betrachten ein klassisches Glücksradspiel, bei dem der Pfeil am Ende auf eines von 54 Felder zeigt.

a) Gib die Ereignismenge  $\Omega$  für einen einzelnen Spin an.

$$\Omega = \{\text{Sa, Am, To, Op, Sm, Ru, Di}\}$$

b) Wie hoch ist die Wahrscheinlichkeit, dass der Pfeil auf einen Topas zeigt?

$$P(\text{Topas}) = \frac{8}{54} = 0.148 = 14.8\%$$

c) Wie hoch ist die Wahrscheinlichkeit, dass der Pfeil auf einen Rubin oder einen Diamanten zeigt? Definiere dazu auch ein entsprechendes Ereignis!

$$A = \{\text{Ru, Di}\}$$

$$P(A) = P(\text{Ru}) + P(\text{Di}) = \frac{1}{54} + \frac{1}{54} = 3.7\%$$

d) Wie hoch ist die Wahrscheinlichkeit, dass dieses Ereignis zweimal in Folge eintritt?

Ereignis  $A_i = \text{Ru oder Di im Spin Nr. } i$

$$P(A_1 \cap A_2) = P(A_1) \cdot P(A_2) = \frac{2}{54} \cdot \frac{2}{54} = 0.137\%$$

e) Wie hoch ist die Wahrscheinlichkeit, dass dieses Ereignis zehnmal in Folge NICHT eintritt?

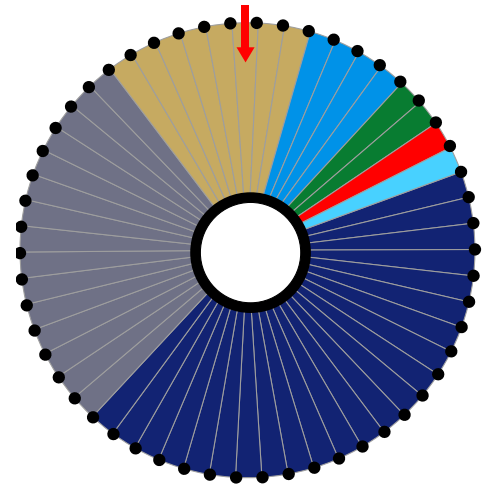
$$P(\overline{A_i}) = 1 - P(A_i) = 96.3\%$$

$$P(\overline{A_1} \cap \dots \cap \overline{A_{10}}) = P(\overline{A_i})^{10} = 68.56\%$$

f) Wie hoch ist die Wahrscheinlichkeit, dass bei 5 Spins mindestens ein Amethyst auftritt?

Ereignis  $B_i = \text{Amethyst im Spin Nr. } i$

$$P(B_1 \cup \dots \cup B_5) = 1 - P(\overline{A_1}) \cdot \dots \cdot P(\overline{A_5}) = 1 - 0.722^5 = 80.4\%$$



23x Saphir

2x Smaragd

15x Amethyst

1x Rubin

8x Topas

1x Diamant

4x Opal

$\Sigma = 54$  Felder

## Aufgabe 14 Permutationen und Binomialkoeffizienten

a) Sie bestücken ein frisch montiertes Regal mit acht Büchern. Wie viele mögliche Anordnungen gibt es, wenn Sie genau acht Bücher dabei zur Auswahl haben?



$$n! = \prod_{i=1}^8 i = 1 \cdot 2 \cdot \dots \cdot 8 = 40320$$

b) Wie viele Auswahlmöglichkeiten gibt es, wenn Sie zwölf Bücher zur Auswahl haben aus denen Sie acht für das Regal auswählen? Wie viele verschiedene Anordnungen an Büchern gibt es insgesamt?

$$\text{Auswahlmögl.: } \binom{12}{8} = \frac{12!}{8! (12-8)!} = \frac{12!}{8! \cdot 4!} = 495$$

$$\text{Anordnungen: } \binom{12}{8} 8! = \frac{12!}{\cancel{8!} (12-8)!} \cancel{8!} = 19\,958\,400$$

c) Ein Bogenschütze trifft ein Ziel mit einer Wahrscheinlichkeit von 80%. Berechnen Sie die Wahrscheinlichkeiten, dass er von 3 Schüssen genau 0, 1, 2 bzw. 3 Schüsse trifft.

$$P_k = \binom{n}{k} P_T^k (1-P_T)^{n-k}$$

$$P_0 = \binom{3}{0} \cdot 0.8^0 \cdot 0.2^3 = 0.8\%$$

$$P_1 = \binom{3}{1} \cdot 0.8^1 \cdot 0.2^2 = 9.6\%$$

$$P_2 = \binom{3}{2} \cdot 0.8^2 \cdot 0.2^1 = 38.4\%$$

$$P_3 = \binom{3}{3} \cdot 0.8^3 \cdot 0.2^0 = 51.2\%$$

d) In einer Packung „Das gute Himbeermüsli“ vom Discounter Brutto sollte genau eine Himbeere enthalten sein. Durch einen Produktionsfehler fehlt diese Himbeere jedoch in jeder 100-sten Packung.

Die Verbraucherschutzzentrale führt eine Stichprobe mit 20 Packungen durch. Findet Sie mindestens zwei Packungen ohne Himbeere, kann Sie den Discounter belangen.

Wie hoch ist die Wahrscheinlichkeit, dass ihr dies gelingt?

$$P_F = \frac{1}{100} = 1\%$$

$$\left. \begin{aligned} P_0 &= \binom{20}{0} \cdot 0.01^0 \cdot 0.99^{20} = 81.79\% \\ P_1 &= \binom{20}{1} \cdot 0.01^1 \cdot 0.99^{19} = 16.52\% \end{aligned} \right\} 1 - P_0 - P_1 = 1.69\%$$



### Aufgabe 15 Urnenmodelle I

Wir betrachten das Glücksspiel KENO-10. Auf dem KENO-Schein kreuzen wir 10 von 70 Zahlen an. Gezogen werden insgesamt 20 Gewinnzahlen.

Bei dem KENO-10 Spiel haben wir Ziehen ohne Zurücklegen und mit irrelevanter Reihenfolge.

Die Anzahl möglicher Kombinationen an Gewinnzahlen ist:

$$\binom{n}{k} = \binom{70}{20} = \frac{70!}{20!(70-20)!} = 161.884.603.662.65$$

a) Wie hoch ist die Wahrscheinlichkeit, dass alle 10 getippten Zahlen eine Gewinnzahl sind?

Die Anzahl an Ziehungen, bei denen 10 Zahlen richtig sind ist:

$$\binom{10}{10} \binom{60}{10} = 75\,394\,027\,566$$

Die Wahrscheinlichkeit für 10 richtige (Ereignis A) ist damit:

$$P(A) = \frac{75394027566}{161884603662657876} = 0.000046572\%$$

Für die nächsten beiden Teilaufgaben definieren wir Ereignis B als genau 9 Gewinnzahlen und Ereignis C als keine einzige Gewinnzahl.



b) Wie hoch ist die Wahrscheinlichkeit, dass genau 9 von unseren 10 getippten Zahlen Gewinnzahlen sind?

$$P(B) = \frac{\binom{10}{9} \binom{60}{11}}{\binom{70}{20}} = 0.00211\%$$

c) Wie hoch ist die Wahrscheinlichkeit, dass keine einzige der 10 getippten Zahlen Gewinnzahlen sind?

$$P(C) = \frac{\binom{10}{0} \binom{60}{20}}{\binom{70}{20}} = 2.58940\%$$

### Aufgabe 16 Urnenmodelle II

Betrachten Sie ein elektronisches Zahlenschloss, bei dem wir mit den Ziffern 0-9 einen fünfstelligen Code eingeben müssen.

a) Wie hoch ist die Wahrscheinlichkeit, dass wir mit einem Versuch genau den richtigen Code eingeben?

$$P = \frac{1}{10^5} = 0.001\%$$

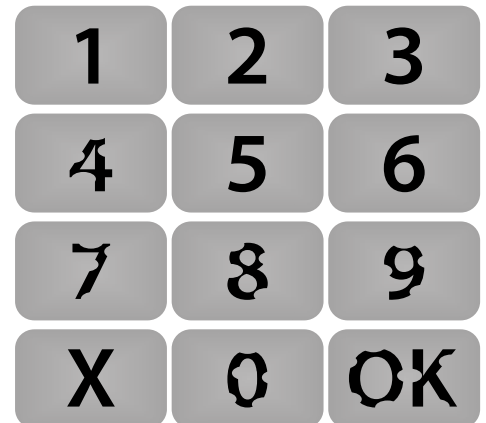
b) Wie viele verschiedene Kombinationen müssen wir ausprobieren, damit wir zu 50% Wahrscheinlichkeit den richtigen Code finden?

Ziemlich genau 50000

c) Wie viele sind es, wenn wir uns nicht merken welche Kombinationen wir bereits erfolglos probiert haben und bei jedem Versuch einfach zufällig 5 Tasten drücken?

Definiere das Ereignis  $A_i$  als ein Öffnen der Türe in Versuch Nr.  $i$  und rechne über die Gegenwahrscheinlichkeiten.

2 1 9 \_ \_



$$\begin{aligned}
P(A_1 \cup \dots \cup A_n) &= 1 - P(\bar{A}_1) \cdot \dots \cdot P(\bar{A}_n) \stackrel{!}{=} 0.50 \\
&\Leftrightarrow 1 - 0.99999^n = 0.50 \\
&\Leftrightarrow 0.99999^n = 0.50 \\
&\Leftrightarrow n = \log_{0.99999}(0.50) = 69314.4
\end{aligned}$$

d) Der Programmierer des Schlosses hat einen Fehler gemacht. Es öffnet die Türe, wenn von den 5 Ziffern mindestens 4 richtig sind. Wie hoch ist die Wahrscheinlichkeit die Tür im ersten Versuch aufzubekommen?

Die Wahrscheinlichkeit von mindestens 4 aus 5 richtigen Ziffern ist gegeben durch:

$$\begin{aligned}
P_k &= \binom{n}{k} P_R^k (1-P_R)^{n-k} \\
P_4 &= \binom{5}{4} \cdot 0.1^4 \cdot 0.9^1 = 0.045\% \\
P_5 &= \binom{5}{5} \cdot 0.1^5 \cdot 0.2^0 = 0.001\% \\
\hline
&0.046\%
\end{aligned}$$

e) Der Kundendienst behebt den Fehler der letzten Teilaufgabe, baut dabei aber versehentlich einen neuen Fehler ein: Jetzt spielt die Reihenfolge der Ziffern keine Rolle mehr, d.h. wenn 10125 der richtige Code wäre, würden Sie auch mit 25110 die Tür öffnen können. Wie hoch ist die Wahrscheinlichkeit die Tür im ersten Versuch aufzubekommen?

Aus Ziehen mit Zurücklegen und mit relevanter Reihenfolge wird jetzt Ziehen mit Zurücklegen, jedoch irrelevanter Reihenfolge. Statt 100000 möglicher Kombinationen gibt es nur noch:

$$\begin{aligned}
\binom{n+k-1}{k} &= \binom{10+5-1}{5} = \frac{14!}{5!(14-5)!} = 2002 \\
P &= \frac{1}{2002} = 0.04995\%
\end{aligned}$$

f) Nachdem das Schloss endlich fehlerfrei programmiert ist, tut sich ein anderes Problem auf: Durch die Benutzung der Eingabetasten sind die Zahlen aus denen der Code besteht offensichtlich: 0,4,7,8 und 9. Wie hoch ist die Wahrscheinlichkeit die Tür im ersten Versuch aufzubekommen?

Da es fünf verschiedene Zahlen für 5 Stellen sind:  $1/5! = 1/120 = 0.833\%$

## Aufgabe 17 Urnenmodelle III

Bei der Pokervariante Five-Card-Draw erhält jeder Spieler fünf Handkarten aus einem 52-er Kartendeck.

Bei der Aufgabe handelt es sich um Ziehen ohne Zurücklegen und mit irrelevanter Reihenfolge. Die Anzahl unterschiedlicher Pokerhände wäre bei dieser Variante:

$$\binom{52}{5} = \frac{52!}{5! (52-5)!} = \frac{52!}{5! \cdot 47!} = 2598960$$

Wir definieren die im folgenden untersuchten Kombinationen als Ereignisse A, B und C.

♥ 2	♣ 2	♦ 2	♠ 2
♥ 3	♣ 3	♦ 3	♠ 3
♥ 4	♣ 4	♦ 4	♠ 4
♥ 5	♣ 5	♦ 5	♠ 5
♥ 6	♣ 6	♦ 6	♠ 6
♥ 7	♣ 7	♦ 7	♠ 7
♥ 8	♣ 8	♦ 8	♠ 8
♥ 9	♣ 9	♦ 9	♠ 9
♥ 10	♣ 10	♦ 10	♠ 10
♥ J	♣ J	♦ J	♠ J
♥ Q	♣ Q	♦ Q	♠ Q
♥ K	♣ K	♦ K	♠ K
♥ A	♣ A	♦ A	♠ A

a) Wie hoch ist die Wahrscheinlichkeit eines Flushes, d. h. alle fünf Karten mit demselben Symbol?

Beim Flush benötigen wir 5 Karten mit demselben Symbol. Da es vier verschiedene Symbole gibt und es jeweils 13 Karten mit einem bestimmten Symbol gibt, gilt:

$$P(A) = \frac{4 \cdot \binom{13}{5}}{2598960} = \frac{4 \cdot 13!}{5! (13-5)! \cdot 2598960} = 0.198\%$$

b) Wie hoch ist die Wahrscheinlichkeit eines Vierlings, d. h. vier Karten mit demselben Wert?

Beim Vierling benötigen wir 4 Karten mit demselben Wert. Da es 13 verschiedene Werte gibt, können wir einen davon wählen. Die fünfte Karte ist eine der 48 übrigen Karten.

$$P(B) = \frac{13 \cdot \binom{48}{1}}{2598960} = \frac{13 \cdot 48!}{1! (48-1)! \cdot 2598960} = 0.024\%$$

c) Wie hoch ist die Wahrscheinlichkeit eines Full House, d. h. drei Karten mit demselben Wert und zwei Karten mit einem anderen gleichen Wert?

Beim Full House wählen wir einen Wert aus 13 für den Drilling und einen Wert aus den verbleibenden 12 für das Paar. Beim Drilling benötigen wir 3 von 4 passenden Karten, beim Zwilling 2 von 4.

$$P(C) = \frac{13 \binom{4}{3} \cdot 12 \binom{4}{2}}{2598960} = 0.144\%$$



### Aufgabe 18 Urnenmodelle IV

Betrachten Sie eine Bar mit 20 verschiedenen Spirituosen, Säften und Sirupen, darunter auch die rechts auf dem Rezept genannten.

Bei der Aufgabe handelt es sich um Ziehen mit Zurücklegen und mit irrelevanter Reihenfolge. Die Anzahl unterschiedlicher Mischungen ist:

$$\binom{n+k-1}{k} = \binom{20+8-1}{8} = \frac{27!}{8!(27-8)!} = 2220075$$

Allerdings ist nicht jede Mischung davon gleich wahrscheinlich. Als Basis für die folgenden Rechnungen nehmen wir daher die folgende Anzahl an Ereignissen an:

$$20^8 = 25.600.000.000$$

#### Long Island Iced Tea Rezept

20ml Tequilla  
20ml Vodka  
20ml White Rum  
20ml Cointreau  
20ml Gin  
40ml Zitronensaft  
20ml Zuckersirup

Zutaten gut mischen und im Glas mit Cola auffüllen

a) Sie nehmen einen 20ml Messbecher und füllen ihn 8 Mal mit einer zufällig ausgewählten Flüssigkeit. Wie hoch ist die Wahrscheinlichkeit, dass sie zufällig einen Long Island Iced Tea dabei mischen?

$$p = \frac{\frac{8!}{2!}}{25.600.000.000} = 0.00007875\%$$

Es gibt genau eine richtige Rezeptur, wobei es keine Rolle spielt in welcher Reihenfolge wir die Zutaten hinzufügen. Wären es 8 verschiedene Zutaten gäbe es 8! verschiedene Reihenfolgen. Dadurch, dass wir eine Zutat (Zitronensaft) doppelt benötigen gibt es jedoch nur halb so viele Reihenfolgen.

### Aufgabe 19 Bedingte Wahrscheinlichkeit I

Ein Angler geht jeden Sonntagabend zum Angeln. Bei guten Wetterbedingungen fängt er zu 75% einen Fisch, bei schlechten Wetterbedingungen nur zu 40%. Die Wahrscheinlichkeit für ein geeignetes Wetter liegt bei 60%.

a) Definiere die Zufallsereignisse und schreibe die gegebenen Wahrscheinlichkeiten als Gleichungen.

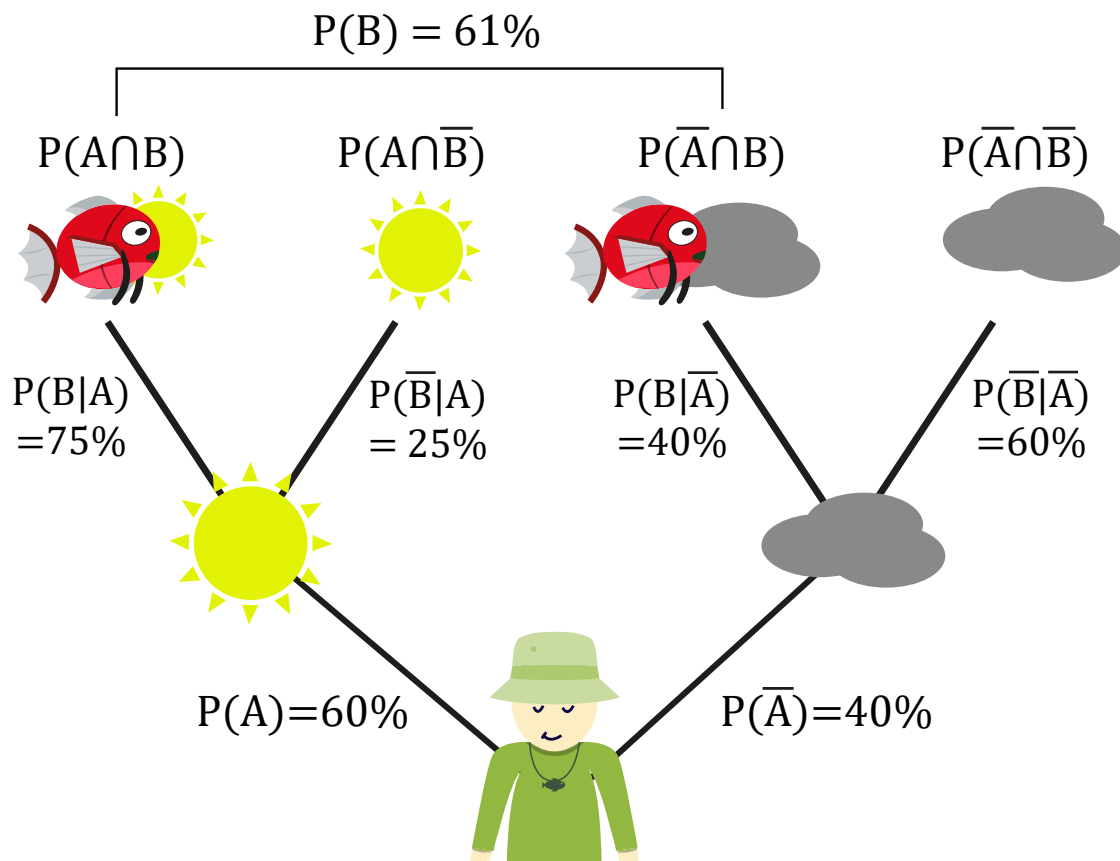
Ereignis A = Geeignetes Wetter mit  $P(A) = 0.60$

Ereignis B = Fisch gefangen mit  $P(B | A) = 0.75$  und  $P(B | \bar{A}) = 0.40$

b) Berechne die Wahrscheinlichkeit, dass der Angler am nächsten Sonntag einen Fisch fängt.

$$P(B) = P(A) P(B | A) + P(\bar{A}) P(B | \bar{A}) = 0.6 \cdot 0.75 + 0.4 \cdot 0.4 = 61\%$$

c) Stelle alle Ereignisse und Wahrscheinlichkeiten in Form eines Wahrscheinlichkeitsbaums dar.



## Aufgabe 20 Bedingte Wahrscheinlichkeit II

Ein Agrarhändler verkauft ein „Trächtigkeitsprüfgerät“ für Rinder, dass ein trächtiges Rind mit 95% Sicherheit feststellen kann und ein nicht-trächtiges Rind zu 99% als nicht trächtig einstuft. Der Test soll verhindern, dass trächtige Rinder geschlachtet werden, denn laut Statistik sind 5% aller geschlachteter Rinder trächtig gewesen.

a) Definiere die Zufallsereignisse und schreibe die gegebenen Wahrscheinlichkeiten als Gleichungen.

Ereignis A = Eine trächtige Kuh mit  $P(A) = 0.05$

Ereignis B = Ein positiver Test mit  $P(B | A) = 0.95$  und  $P(B | \bar{A}) = 0.01$

a) Wie hoch ist die Wahrscheinlichkeit, dass ein als trächtig angezeigtes Rind auch trächtig ist?

Als Erstes benötigen wir die Wahrscheinlichkeit eines positiven Tests:

$$\begin{aligned} P(B) &= P(B|A) P(A) + P(B|\bar{A}) P(\bar{A}) \\ &= 0.95 \cdot 0.05 + 0.01 \cdot 0.95 \\ &= 0.057 \end{aligned}$$

Dann können wir den Satz von Bayes anwenden!

$$\begin{aligned} P(A | B) &= \frac{P(B | A) \cdot P(A)}{P(B)} \\ &= \frac{0.95 \cdot 0.05}{0.057} = 83.33\% \end{aligned}$$

b) Wie hoch ist die Wahrscheinlichkeit trotz eines negativen Tests ein trächtiges Rind zu schlachten?

$$\begin{aligned} P(A | \bar{B}) &= \frac{P(\bar{B} | A) \cdot P(A)}{P(\bar{B})} \\ &= \frac{0.05 \cdot 0.05}{0.943} = 0.265\% \end{aligned}$$

### Aufgabe 21 Bedingte Wahrscheinlichkeit III

„6 von 10 tödlich verunglückten Autofahrern trugen einen Sicherheitsgurt“ - Um diese Headline besser einschätzen zu können recherchieren sie, dass 99% aller Autofahrer einen Sicherheitsgurt tragen. Eine Wahrscheinlichkeit für einen tödlichen Unfall können Sie jedoch nicht finden.

a) Definiere die beiden Zufallsereignisse und stelle die gegebenen Wahrscheinlichkeiten formal dar.

Ereignis A = Autofahrer trägt Gurt  $P(A) = 0.99$

Ereignis B = Ein tödlicher Unfall mit  $P(A|B) = 0.60$

b) Versuche mit den gegebenen Wahrscheinlichkeiten herauszufinden, ob Sicherheitsgurte positiv, neutral oder negativ für die Sicherheit des Fahrers sind.

Wir können zwar nicht die konkreten Wahrscheinlichkeiten ausrechnen mit der wir mit/ohne Gurt tödlich verunglücken, aber wir können ein Gefahrenverhältnis ausrechnen! Mit dem Sicherheitsgurt ist die Gefahr etwa 50 mal kleiner ...

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

$$P(\bar{A} | B) = \frac{P(B | \bar{A}) \cdot P(\bar{A})}{P(B)} \iff P(B | \bar{A}) = \frac{P(\bar{A} | B) \cdot P(B)}{P(\bar{A})}$$

$$\begin{aligned} \frac{P(B | A)}{P(B | \bar{A})} &= \frac{\frac{P(A | B) \cdot P(B)}{P(A)}}{\frac{P(\bar{A} | B) \cdot P(B)}{P(\bar{A})}} = \frac{P(A | B) \cdot P(\bar{A})}{P(\bar{A} | B) \cdot P(A)} \\ &= \frac{0.6 \cdot 0.01}{0.4 \cdot 0.99} = 0.015 \end{aligned}$$

## Aufgabe 22 Zufallsvariablen I

Betrachten Sie die rechts gezeigte folgende Verteilungsfunktion der Zufallsvariable X!

a) Geben Sie die Bedeutung der drei markierten Punkte verbal an.

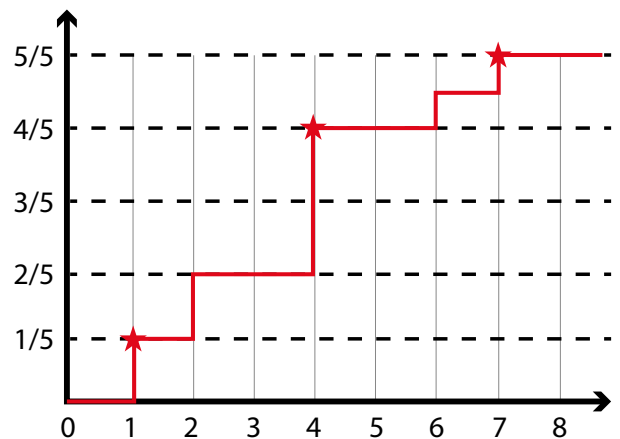
Die Wahrscheinlichkeit, dass  $X \leq 1$  ist, beträgt 20%. Da die Verteilungsfunktion für alle Zahlen unter 1 den Wert 0 annimmt, wissen wir außerdem, dass 1 der kleinstmögliche Wert von X ist.

Die Wahrscheinlichkeit, dass  $X \leq 4$  ist, beträgt 80%.

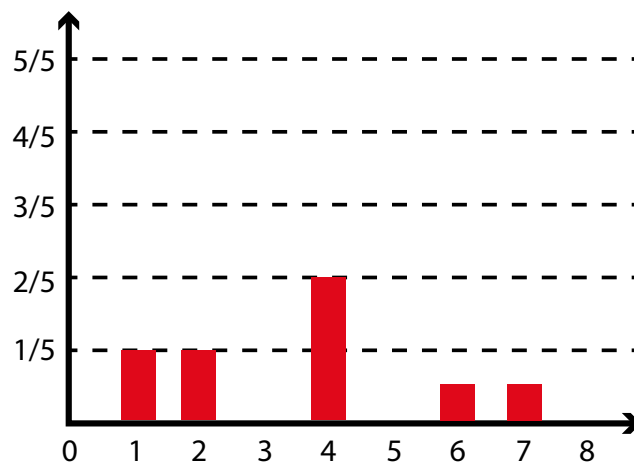
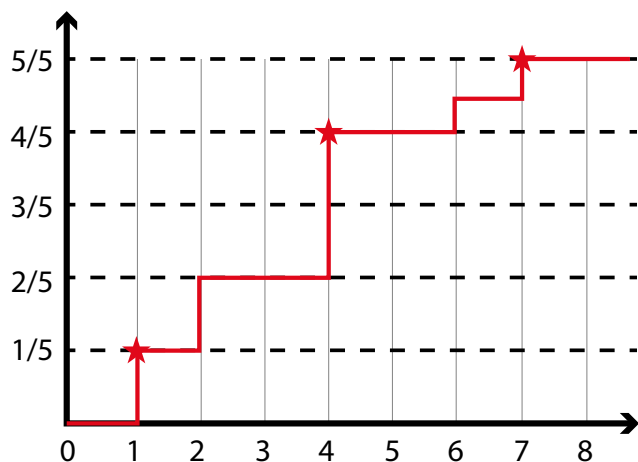
Die Wahrscheinlichkeit, dass  $X \leq 7$  ist, beträgt 100%. Da die Verteilungsfunktion an dieser Stelle zum ersten Mal den Wert 1 annimmt, wissen wir außerdem, dass 7 der größtmögliche Wert von X ist.

b) Wie hoch ist die Wahrscheinlichkeit, dass X mindestens 2.5 ist?

Die Wahrscheinlichkeit, dass  $X \leq 2.5$  ist, beträgt 40%. Gefragt ist jedoch nach dem Gegenteil davon: X soll mindestens 2.5 betragen! Die Wahrscheinlichkeit, dass  $X \geq 2.5$  ist  $1 - 2/5 = 60\%$



c) Zeichne eine zur Verteilungsfunktion passende Dichtefunktion!

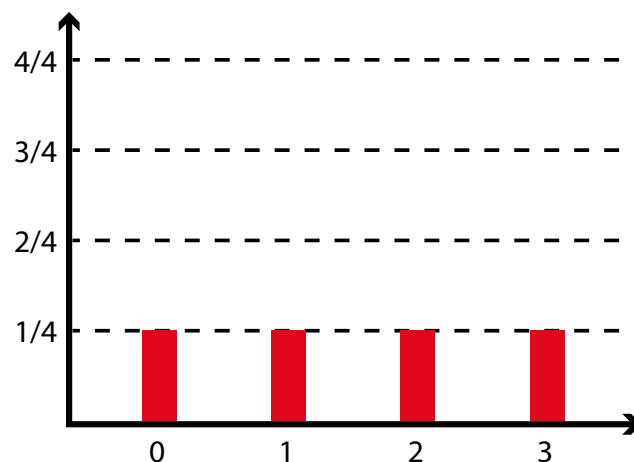
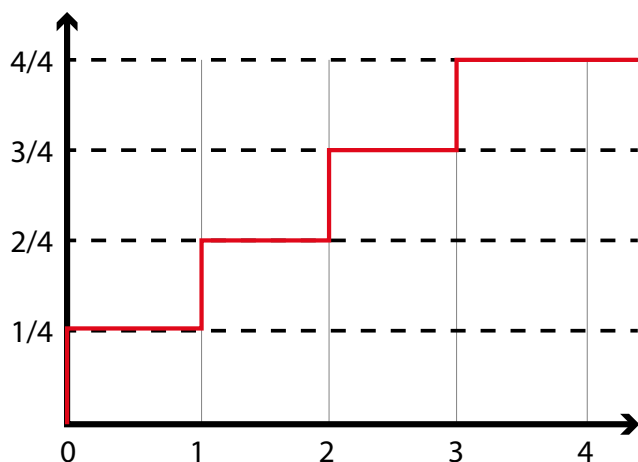


d) Berechne den Erwartungswert von X.

$$\begin{aligned}
 E(X) &= \sum_{x \in E} x \cdot f(x) \\
 &= 1 \cdot \frac{2}{10} + 2 \cdot \frac{2}{10} + 4 \cdot \frac{4}{10} + 6 \cdot \frac{1}{10} + 7 \cdot \frac{1}{10} \\
 &= 0.2 + 0.4 + 1.6 + 0.6 + 0.7 = 3.5
 \end{aligned}$$

### Aufgabe 23 Zufallsvariablen II

Wir betrachten einen vierseitigen Würfel, der mit einer Wahrscheinlichkeit von jeweils 1/4 die Augenzahlen 0, 1, 2 und 3 zeigt. Die Zufallsvariable X entspricht der gewürfelten Augenzahl.



b) Berechne den Erwartungswert und die Varianz von X.

$$E(X) = \sum_{x \in E} x \cdot f(x)$$

$$= 0 \cdot \frac{1}{4} + 1 \cdot \frac{1}{4} + 2 \cdot \frac{1}{4} + 3 \cdot \frac{1}{4}$$

$$= 0.0 + 0.25 + 0.50 + 0.75 = 1.5$$

$$\text{Var}(X) = \sum_{x \in E} [x - E(X)]^2 \cdot f(x)$$

$$= (-1.5)^2 \frac{1}{4} + (-0.5)^2 \frac{1}{4} + (0.5)^2 \frac{1}{4} + 1.5^2 \frac{1}{4}$$

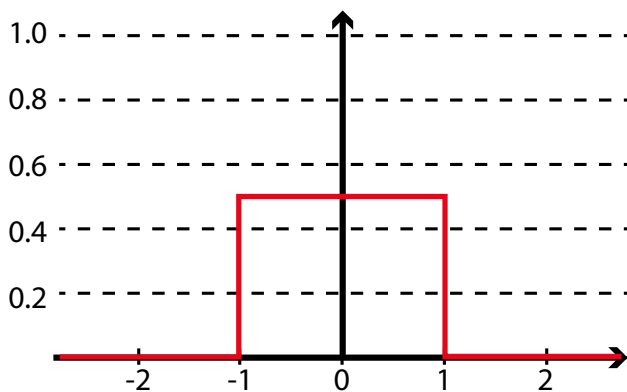
$$= 1.25$$

### Aufgabe 24 Zufallsvariablen III

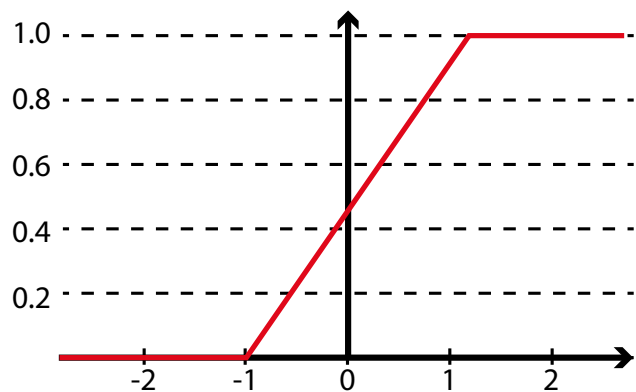
Zwei Zufallsgeneratoren erzeugen gleichverteilte Zahlen auf den Intervallen -1 bis 1 und 10 bis 20. Die Zufallsvariable X (-1 bis 1) und Y (10 bis 20) sind die jeweils generierten Zahlen.

a) Stelle die Verteilung von X und Y als Verteilungsfunktion und als Dichtefunktion dar.

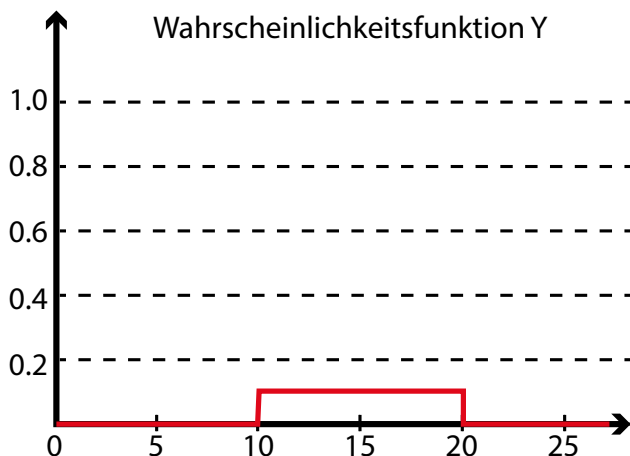
Wahrscheinlichkeitsfunktion X



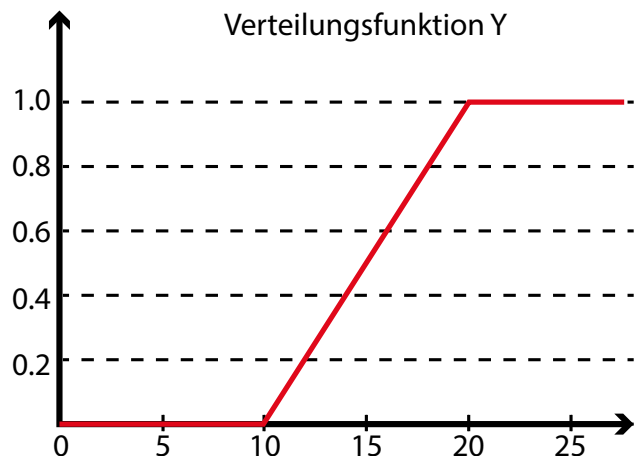
Verteilungsfunktion X



Wahrscheinlichkeitsfunktion Y



Verteilungsfunktion Y



b) Berechne den Erwartungswert von X, Y und X+Y.

$$E(X) = \int_{-\infty}^{\infty} x \cdot \varphi(x) dx = \int_{-1}^1 0.5x dx = \left[ 0.25x^2 \right]_{-1}^1 = 0.25 \cdot 1^2 - 0.25 \cdot (-1)^2 = 0$$

$$E(Y) = \int_{-\infty}^{\infty} y \cdot \varphi(y) dy = \int_{10}^{20} 0.1 dy = \left[ 0.05x^2 \right]_{10}^{20} = 0.05 \cdot 20^2 - 0.05 \cdot 10^2 = 15$$

$$E(X+Y) = E(X) + E(Y) = 0 + 15 = 15$$

c) Berechne die Varianz von X und Y.

$$\begin{aligned} \text{Var}(X) &= \int_{-\infty}^{\infty} [x - E(X)]^2 \cdot \varphi(x) dx = \int_{-1}^1 (x - 0)^2 \cdot 0.5 dx \\ &= \int_{-1}^1 0.5x^2 dx = \left[ \frac{1}{6} x^3 \right]_{-1}^1 = \frac{1}{3} \end{aligned}$$

$$\begin{aligned} \text{Var}(Y) &= \int_{10}^{20} [y - E(Y)]^2 \cdot \varphi(y) dy = \int_{10}^{20} (y-15)^2 \cdot 0.1 dy \\ &= \int_{10}^{20} 0.1y^2 - 3y + 22.5 dy = \left[ \frac{1}{30} y^3 - 1.5y^2 + 22.5y \right]_{10}^{20} \\ &= 8.333 \end{aligned}$$

## Aufgabe 25 Zufallsvariablen IV

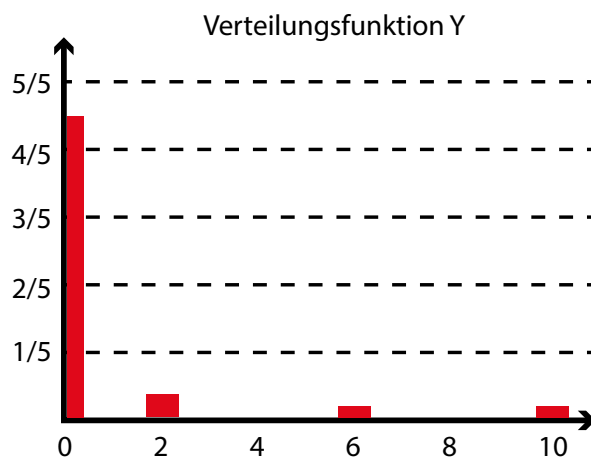
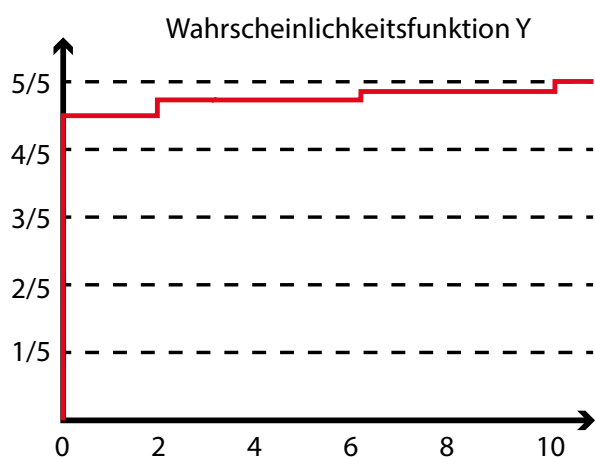
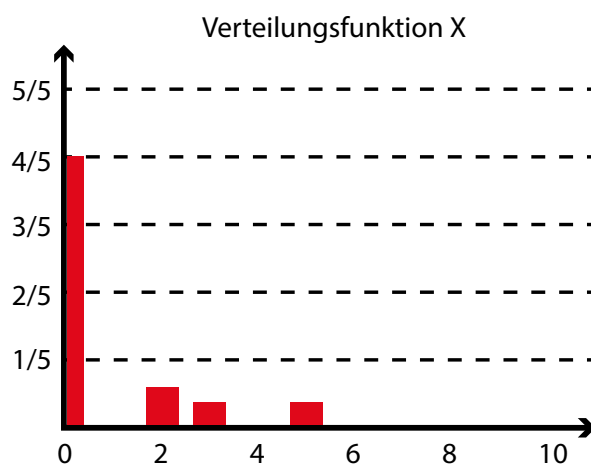
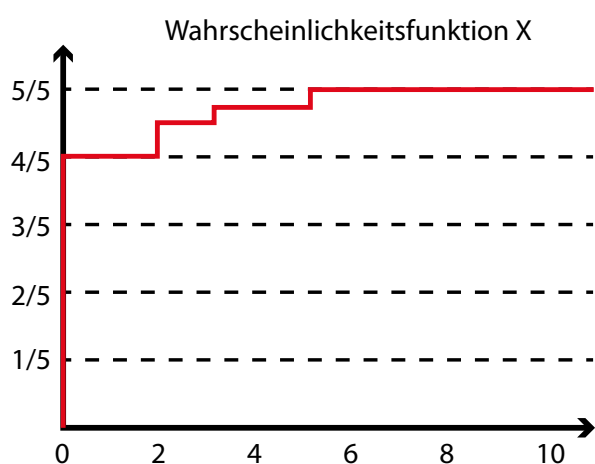
Wir betrachten zwei verschiedene Anbieter von Rubbellosen: King-Cash und BigWin. Die Tabelle rechts gibt die Häufigkeit von Nieten und Gewinnlosen an. Beide Anbieter verlangen 1€ pro Los.

Die Zufallsvariable  $X$  beschreibt die Auszahlung eines KingCash Loses, die Zufallsvariable  $Y$  beschreibt dagegen die Auszahlung eines BigWin Loses.

Von 1000 Losen sind...

KingCash		BigWin	
800x	Nieten	900x	Nieten
100x	2€	50x	2€
50x	3€	25x	6€
50x	5€	25x	10€

a) Stelle die Verteilung von  $X$  und  $Y$  als Verteilungsfunktion und als Wahrscheinlichkeitsfunktion dar.





b) Berechne den Erwartungswert und die Varianz von X und Y

$$E(X) = \sum_{x \in E} x \cdot f(x)$$

$$= 2 \cdot 0.1 + 3 \cdot 0.05 + 5 \cdot 0.05$$

$$= 0.6$$

$$\text{Var}(X) = \sum_{x \in E} [x - E(X)]^2 \cdot f(x)$$

$$= (-0.6)^2 \cdot 0.8 + (1.4)^2 \cdot 0.1 + (2.4)^2 \cdot 0.05 + (4.4)^2 \cdot 0.05$$

$$= 1.74$$

$$E(Y) = \sum_{y \in E} y \cdot f(y)$$

$$= 2 \cdot 0.05 + 6 \cdot 0.025 + 10 \cdot 0.025$$

$$= 0.5$$

$$\text{Var}(Y) = \sum_{y \in E} [y - E(Y)]^2 \cdot f(y)$$

$$= (-0.5)^2 \cdot 0.9 + (1.5)^2 \cdot 0.05 + (5.5)^2 \cdot 0.025 + (9.5)^2 \cdot 0.025$$

$$= 3.35$$

c) Berechne den Erwartungswert und die Varianz von  $5X+5Y$ , d. h. die Auszahlung aus dem Kauf von jeweils fünf KingCash Losen und fünf BigWin Losen.

Da die Auszahlung der gekauften Lose statistisch unabhängig sind erhalten wir einen Erwartungswert von 5.50€ und eine Varianz von 25.45

### Aufgabe 26 Z-Test

Die beiden Anbieter KingCash und BigWin werden einer Untersuchung unterzogen, da Zweifel bestehen, ob die Anzahl an Nieten nicht deutlich höher ist als in der Tabelle rechts oben angegeben.

Die zuständige Behörde kauft von jedem Anbieter 500 Lose und findet 415 Nieten bei KingCash und 470 Nieten bei BigWin.

Verwende einen einseitigen Z-Test, um einschätzen zu können, ob sich die beiden Hersteller auf juristische Konsequenzen einstellen sollten. Stelle dazu geeignete Hypothesenpaare auf und prüfe, ob die zu große Anzahl an Nieten in der Stichprobe signifikant sind oder plausibel als Zufall eingeordnet werden könnten.

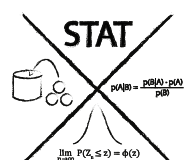
Gegeben der Tabelle sollten wir  $X=400$  Nieten bei KingCash und  $Y=450$  Nieten bei BigWin finden.

Wir definieren die Zufallsvariablen  $x_i$  und  $y_i$ . Diese ZVen nehmen den Wert 1 an, wenn das i-te Los eine Niete ist und den Wert 0, wenn das i-te Los ein beliebiger Gewinn ist.

Die Erwartungswerte sind dann  $E[x_i]=0.8$  und  $E[y_i]=0.9$ . Bei den Varianzen und Standardabweichungen müssen wir die in der Kombinatorik gelernten Formeln anwenden.

Von 1000 Losen sind...

KingCash	BigWin
800x Nieten	900x Nieten
100x 2€	50x 2€
50x 3€	25x 6€
50x 5€	25x 10€



$$\begin{aligned}\text{Var}(x_i) &= \sum_{x \in E} [x_i - E(x_i)]^2 \cdot f(x_i) & \sigma_x &= \sqrt{0.16} = 0.4 \\ &= (-0.8)^2 \cdot 0.2 + (0.2)^2 \cdot 0.8 = 0.16\end{aligned}$$

$$\begin{aligned}\text{Var}(y_i) &= \sum_{y \in E} [y_i - E(y_i)]^2 \cdot f(y_i) & \sigma_y &= \sqrt{0.09} = 0.3 \\ &= (-0.9)^2 \cdot 0.1 + (0.1)^2 \cdot 0.9 = 0.09\end{aligned}$$

Die Standardabweichungen für 500 gekaufte Lose sind:

$$\begin{aligned}\text{Var}(X) &= \sum_{i \in 1}^{500} \text{Var}(x_i) = 500 \cdot 0.16 = 80 & \sigma_x &= \sqrt{80} = 8.95 \\ \text{Var}(Y) &= \sum_{i \in 1}^{500} \text{Var}(y_i) = 500 \cdot 0.09 = 45 & \sigma_y &= \sqrt{45} = 6.71\end{aligned}$$

Bei KingCash erhalten wir 415 statt den 400 erwarteten Nieten. Die Abweichung beträgt 1.675 Standardabweichungen. Unter Gültigkeit der Nullhypothese ...

$H_{0_A}$  Die Häufigkeit von Nieten beim Anbieter KingCash beträgt wie angegeben 80%.

$H_{1_A}$  Die Häufigkeit von Nieten beim Anbieter KingCash ist höher als angegeben.

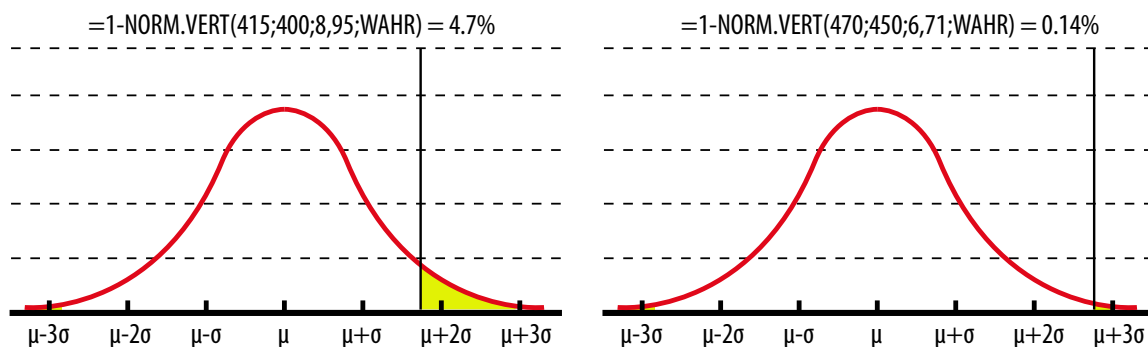
... ist das eher unwahrscheinlich. In diesem Fall erhalten wir bei 500 Losen mit einer Wahrscheinlichkeit von 4.7% genau 415 oder noch mehr Nieten.

Bei BigWin erhalten wir 470 statt den 450 erwarteten Nieten. Die Abweichung beträgt 2.98 Standardabweichungen. Unter Gültigkeit der Nullhypothese ...

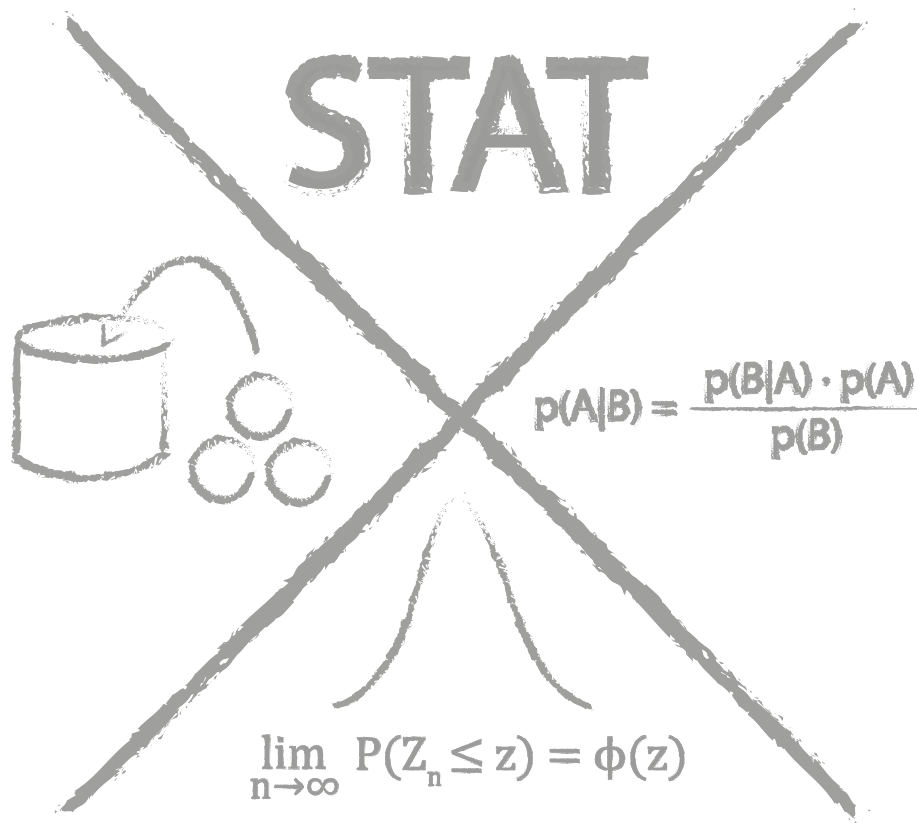
$H_{0_B}$  Die Häufigkeit von Nieten beim Anbieter BigWin beträgt wie angegeben 90%.

$H_{1_B}$  Die Häufigkeit von Nieten beim Anbieter BigWin ist höher als angegeben.

... ist das sehr unwahrscheinlich. In diesem Fall erhalten wir bei 500 Losen mit einer Wahrscheinlichkeit von 0.14% genau 470 oder noch mehr Nieten.







Bereitgestellt durch das Zentrum für Ökonomik (ZÖK)  
der DHBW Ravensburg