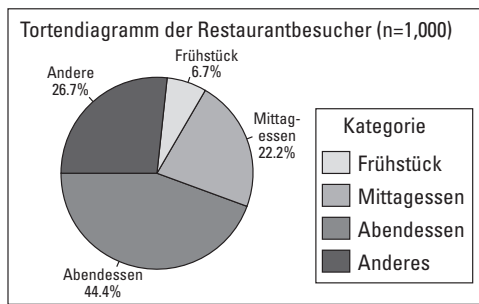


# Statistik für Dummies – Zusatzmaterial

## Aufgabe Kapitel 4

Angenommen, ein Restaurantbesitzer beobachtet, zu welchen Zeiten seine Gäste das Restaurant besuchen: zum Frühstück, Mittagessen, Abendessen oder zu einer anderen Zeit. Einen Monat lang notiert er genau die Besuchszeit von allen seinen Gästen. Insgesamt besuchen in diesem Monat 1.000 Gäste das Restaurant. Das Tortendiagramm gibt die Ergebnisse seiner Beobachtung wieder.



1. Welche Erkenntnisse lassen sich aus diesem Diagramm für den Restaurantbesitzer ablesen?
2. Welches Problem könnte es mit der Kategorie »Anderes« geben? Wie könnte eine solche Untersuchung in der Zukunft verbessert werden?

## Aufgabe Kapitel 5

Angenommen, das Durchschnittsgehalt in einer bestimmten Firma betrage 100.000 €, und der Median aller Gehälter sei 40.000 €.

1. Was können Sie aus diesen beiden Kennzahlen über die Form des Histogramms aller Gehälter der Firma schließen?
2. Welche der beiden Kennzahlen für das Zentrum der Daten ist in diesem Fall geeigneter?
3. Wenn die Firma nun mit der Gewerkschaft in eine Lohnverhandlungsrunde eintritt, wie könnten dann die beiden Parteien die Kennzahlen nutzen, um ihre jeweiligen Standpunkte zu untermauern?

## Aufgabe Kapitel 6

---

Sie werfen eine ungezinkte Münze zehn Mal und notieren dabei, wie häufig die Münze mit dem Kopf nach oben liegen bleibt.

1. Zu wie vielen unterschiedlichen Ergebnissen können die zehn Münzwürfe führen?
2. Wie hoch ist die Wahrscheinlichkeit für jedes der möglichen Ergebnisse?
3. In wie vielen der unterschiedlichen Ergebnisse erscheint genau einmal der Kopf? Wie hoch ist die Wahrscheinlichkeit dafür, dass bei zehn Münzwürfen genau einmal der Kopf (und damit neunmal die Zahl) erscheint?
4. Wie viele Ergebnisse gibt es, in denen kein einziges Mal der Kopf erscheint? Wie hoch ist die Wahrscheinlichkeit, dass die Münze bei zehn Würfeln nie mit dem Kopf nach oben liegen bleibt?
5. Und wie hoch ist die Wahrscheinlichkeit, dass bei zehn Münzwürfen entweder einmal oder gar nicht der Kopf geworfen wird?

## Aufgabe Kapitel 7

---

Martin sitzt bereits seit vier Stunden vor demselben Spielautomaten, den er bereits mit einem Haufen von Münzen gefüttert hat. Er hatte heute noch kein Glück, möchte aber den Spielautomaten gerade jetzt nicht verlassen, da er glaubt, je länger er spielt, desto höher ist die Wahrscheinlichkeit, dass er demnächst gewinnen wird. Hat unser armer Martin recht?

## Aufgabe Kapitel 11

---

Tim geht regelmäßig joggen und benötigt für eine Runde um den See durchschnittlich 8 Minuten, wobei die Rundenzeiten normalverteilt mit einer Standardabweichung von einer Minute variieren. Welche Zeit muss Tim laufen, um in den Top 10% seiner bisher schnellsten Zeiten zu landen?

## Aufgabe Kapitel 12

---

Sie untersuchen 100 Fische aus einem Aquarium A und stellen fest, dass diese im Durchschnitt 55 cm lang sind mit einer Standardabweichung von 10 cm. Eine gleichartige Untersuchung von 100 Fischen aus Aquarium B liefert die gleiche durchschnittliche Länge für die Fische, jedoch hier mit einer Standardabweichung von 20 cm. Wie verhalten sich die Fehlergrenzen für die beiden Stichproben zueinander (unter der Annahme gleicher Konfidenzniveaus)?

## Aufgabe Kapitel 14

Angenommen, ein 95%-Konfidenzintervall für die durchschnittliche Anzahl an Minuten, die Handy-Besitzer im Monat telefonieren, betrage 110 plus/minus 35.

1. Wie hoch ist die Fehlergrenze?
2. Wo liegen die obere und die untere Grenze des Konfidenzintervalls?

## Aufgabe Kapitel 15

1. Angenommen, 73% aller Befragten aus einer Stichprobe von 1.000 Studenten gaben an, ein gebrauchtes Auto zu fahren (die übrigen 27% fahren entweder ein neues oder gar kein Auto). Ermitteln Sie das 80%-Konfidenzintervall für den Anteil aller Studenten, die ein Gebrauchtfahrzeug fahren.
2. Wie groß müsste die Stichprobe sein, um die Fehlergrenze zu halbieren?

## Aufgabe Kapitel 16

Bedeutet ein kleiner p-Wert, dass die Nullhypothese falsch ist? Begründen Sie Ihre Antwort.

## Aufgabe Kapitel 17

Angenommen, Sie haben einen Würfel und möchten überprüfen, ob dieser Würfel tatsächlich fair und unverfälscht ist. Hierzu untersuchen Sie die Häufigkeit, mit der beim Würfeln eine 1 fällt. Formulieren Sie die Null- und die Alternativhypothese für einen Test.

## Aufgabe Kapitel 18

Angenommen, die Ergebnisse einer Wahlforschung zeigen, dass eine Koalition A 48% der Stimmen erreichen wird, während Koalition B auf 52% der Stimmen kommt. Kann man auf Basis dieser Ergebnisse davon ausgehen, dass Koalition B sicher das Rennen macht?

## Aufgabe Kapitel 19

Ein Experiment kommt zu dem Schluss, dass der tägliche Verzehr von einem Ei entgegen den bisherigen Vermutungen von Wissenschaftlern keinen Einfluss auf den

Cholesterinspiegel hat. Das Experiment stützt sich dabei auf junge, gesunde Männer, die sich generell gesund ernähren und wenig Fett zu sich nehmen. Erläutern Sie, was an der Schlussfolgerung der Studie falsch ist.

# Aufgabe Kapitel 20

---

Wie sieht ein Streudiagramm aus, das keinen linearen Zusammenhang zwischen den beiden dargestellten Variablen erkennen lässt?

# Statistik für Dummies – Zusatzmaterial

## Lösung Kapitel 4

Vorsicht ist generell bei Kreisdiagrammen geboten, die eine große Kategorie »Andere« aufweisen.

1. Das Kreisdiagramm zeigt, dass von den drei Mahlzeiten das Abendessen die meisten Gäste lockt. (44% aller Gäste kamen zum Abendessen, 22% zum Mittagessen und 7% zum Frühstück.)
2. Das Kreisdiagramm hat eine Schwäche: Es weist eine sehr große Kategorie »Andere« auf, der 27% aller Gäste des Restaurants zugeordnet sind. Aus dieser Kategorie kann der Restaurantbesitzer keine Anhaltspunkte für die Optimierung seines Geschäfts ableiten, denn es ist nicht zu erkennen, wann genau die 27% der Gäste aus dieser Kategorie das Restaurant besucht haben. Besser wäre es gewesen, die Besuchszeiten der Gäste wären feiner unterteilt worden, beispielsweise mit Hilfe der zusätzlichen Kategorien »Zwischen Frühstück und Mittagessen«, »zum Kaffee«, »am späten Nachmittag« etc.



Wenn Ihnen ein Kreisdiagramm unterkommt, das eine große Kategorie »Sonstiges« enthält, die eventuell sogar größer ist als einige andere Kategorien des Diagramms, sollten Sie aufmerksam werden. Eine große Kategorie »Sonstiges« ist immer ein Hinweis darauf, dass die Daten wohl besser in etwas feinere Kategorien hätten unterteilt werden sollen oder wichtige Kategorien vollständig vergessen wurden.

## Lösung Kapitel 5

Allein aus dem Mittelwert und dem Median lässt sich bereits einiges über die Gehaltsstruktur der Firma lernen:

1. Die Gehaltsverteilung ist rechtsschief. Einige wenige Mitarbeiter in der Firma beziehen sehr hohe Gehälter im Vergleich zu dem Einkommen der übrigen Mitarbeiter. Dies wirkt sich auf den Mittelwert aus, der dadurch entsprechend höher ausfällt, während der Median von den einzelnen hohen Gehältern unbeeinflusst bleibt. Daher liegt der Mittelwert hier deutlich über dem Median.
2. Bei einer schiefen Verteilung ist der Median das geeignetere Maß für das Zentrum der Daten, denn der Median wird, anders als der Mittelwert, nicht von Ausreißern beeinflusst.
3. Die Gewerkschaft würde in den Verhandlungen eher auf den Median verweisen, um das niedrigere Gehaltsniveau der Firma zu dokumentieren, während

## 2 Statistik für Dummies – Zusatzmaterial

die Firmenvertreter eher auf den Mittelwert abstellen würden, um zu belegen, dass die Gehälter auch ohne weitere Lohnerhöhungen bereits ein hohes Durchschnittsniveau erreicht haben.



Bei einem Datensatz mit rechtsschiefer Verteilung der Werte ist der Mittelwert höher als der Median. Sind die Daten dagegen linksschief verteilt, liegt der Mittelwert unterhalb des Medians. Im Falle einer symmetrischen Verteilung sind Mittelwert und Median identisch bzw. liegen sehr nahe beieinander.

## Lösung Kapitel 6

Die Lösung dieser Aufgabe scheint durch die vielen unterschiedlichen Ergebnisse, die bei 10 Münzwürfen eintreten können, sehr aufwendig zu sein – allerdings nur so lange, bis Sie entdecken, dass es neben dem steinigen Weg auch einen einfachen Lösungsweg gibt. Und wenn Sie eine solche Aufgabe in einer Prüfung gestellt bekommen, hat Ihr Professor die verfügbare Zeit zur Lösung der Aufgabe mit Sicherheit so bemessen, dass Sie schon aus Zeitgründen auf die einfache Lösung angewiesen sein werden.

1. Wird eine Münze 10 Mal hintereinander geworfen, kann dies zu  $2 \cdot 2 \cdot 2 \dots \cdot 2$  bzw.  $2^{10} = 1.024$  unterschiedlichen Ergebnissen führen. (Auch als Musterschüler sollten Sie davon absehen, diese Ergebnisse alle sauber aufzuschreiben.)
2. Da die Münze ungezinkt ist, kann jedes Ergebnis mit der gleichen Wahrscheinlichkeit eintreten, also mit einer Wahrscheinlichkeit von  $1 \div 1.024$  bzw. 0,098%.
3. Es gibt 10 verschiedene Möglichkeiten, beim 10-maligen Werfen der Münze genau einmal den Kopf zu erhalten. Ein mögliches Ergebnis lautet zum Beispiel K-Z-Z-Z-Z-Z-Z-Z-Z-Z, ebenso kann der Kopf aber auch beim 2. Wurf, beim 3. Wurf oder an jeder anderen Stelle bis zum 10. Wurf fallen. Die Wahrscheinlichkeit dafür, genau einmal den Kopf zu werfen, beträgt daher  $10 \div 1.024 = 0,0098$  bzw. 0,98%.
4. Wenn bei 10 Würfeln kein einziges Mal der Kopf erscheint, bedeutet dies umgekehrt, dass 10 Mal die Zahl geworfen wird. Dies ist genau bei einem der 1.024 möglichen Ergebnisse der Fall und tritt daher mit einer Wahrscheinlichkeit von 0,00098 bzw. 0,098% ein.
5. Um diese Frage zu beantworten, addieren Sie einfach die Wahrscheinlichkeiten, die Sie in den beiden vorhergehenden Teilaufgaben berechnet haben: Die Wahrscheinlichkeit dafür, dass bei 10 Münzwürfen höchstens einmal der Kopf erscheint, beträgt  $(10 + 1) \div 1.024 = 0,0107$  bzw. 1,07%.

## Lösung Kapitel 7

Nicht wirklich. Auf lange Sicht gesehen sollte Martin davon ausgehen, dass er mit jedem einzelnen Spiel mit großer Wahrscheinlichkeit einen kleinen Geldbetrag verlieren wird.

Natürlich kann es das eine Spiel geben, mit dem er den Super-Gewinn abräumt, die Wahrscheinlichkeit dafür ist aber derart gering, dass er davon ausgehen muss, bis zum Eintreten dieses Super-Gewinns bereits mehr Münzen in die Maschine geworfen zu haben, als er über den Jackpot wieder rausholen wird. Und auch wenn er schon seit Ewigkeiten erfolglos an der Maschine sitzt: Die Wahrscheinlichkeit dafür, dass er mit dem nächsten Spiel den Jackpot knackt, wird dadurch auch nicht größer. Wo bleibt also das ganze Geld, das Martin in die Maschine wirft? Ganz einfach: bei dem Besitzer des Spielautomaten.

## Lösung Kapitel 11

Da nach den »schnellsten« 10% der Rundenzeiten gefragt ist, dürfen nur 10% der bisherigen Zeitmessungen unter dem gesuchten Grenzwert liegen. Benötigt wird also das 10%-Perzentil, das gemäß der Z-Tabelle bei einem Standardwert von ungefähr  $-1,3$  liegt. Anhand der nach  $x$  aufgelösten Z-Formel lässt sich der zugehörige Originalwert berechnen:  $x = -1,3 \cdot 1 + 8 = 6,7$ . Tim muss den See also in nicht mehr als 6,7 Minuten umrundet haben, um in den Top 10% seiner bisherigen Zeiten zu bleiben.

## Lösung Kapitel 12

Die Stichprobe aus Aquarium B weist eine höhere Fehlergrenze auf, da die Standardabweichung größer ist und diese in der Formel zur Berechnung der Fehlergrenze im Zähler des Bruches steht.

## Lösung Kapitel 14

Die Fehlergrenze ist der Wert, den Sie zur Stichprobenkennzahl addieren und subtrahieren, um die obere und untere Grenze des Konfidenzintervalls zu bestimmen.

1. Die Fehlergrenze beträgt hier  $\pm 35$  Minuten.
2. Die untere Grenze des Konfidenzintervalls liegt bei  $110 - 35 = 75$  Minuten die obere Grenze bei  $110 + 35 = 145$  Minuten.

## Lösung Kapitel 15

Die Größe der Fehlergrenze hängt unter anderem von dem Konfidenzniveau und der Stichprobengröße ab.

## 4 Statistik für Dummies – Zusatzmaterial

1. In dieser Aufgabe ist  $n = 1.000$ ,  $\hat{p} = 0,73$  und  $Z^* = 1,28$  (siehe ). Damit ergibt sich ein 80%-Konfidenzintervall von

$$\begin{aligned}\hat{p} \pm Z^* \cdot \sqrt{\frac{\hat{p}(1-\hat{p})}{n}} &= 0,73 \pm 1,28 \cdot \sqrt{\frac{0,73 \cdot (1 - 0,73)}{1.000}} = \\ 0,73 \pm 1,28 \cdot 0,014 &= \\ 0,73 \pm 0,018 &\text{ also } (0,71; 0,75).\end{aligned}$$

Das Konfidenzintervall erstreckt sich also von 71% bis 75%. Mit einer Sicherheit von 80% können Sie davon ausgehen, dass der Anteil aller Studenten, die mit einem Gebrauchtwagen durch die Gegend fahren, zwischen 71% und 75% liegt.

2. Um die Fehlergrenze zu halbieren, müssen Sie die Stichprobengröße verdoppeln. Dies können Sie unmittelbar an der Formel zur Berechnung der Fehlergrenze ablesen: In dieser Formel steht die Wurzel der Stichprobengröße im Nenner des Bruches. Eine Vervielfachung der Stichprobengröße führt also zu einer Verdoppelung des Nenners (denn  $\sqrt{4} = 2$ ) und damit zu einer Halbierung der Fehlergrenze.

## Lösung Kapitel 16

Nein. Ein kleiner p-Wert bedeutet lediglich, dass Sie starke Argumente dafür haben, die Nullhypothese zurückzuweisen. Sie können sich aber niemals sicher sein, ob die Nullhypothese tatsächlich falsch ist. Es besteht immer die Möglichkeit, dass die Nullhypothese wahr ist und nur die Ihnen vorliegenden Stichprobendaten durch zufällige Einflüsse bei der Stichprobenziehung die Nullhypothese als sehr unwahrscheinlich erscheinen lassen.

## Lösung Kapitel 17

Wenn der Würfel unverfälscht ist, wird jede Seite des Würfels bei einem Sechstel aller Würfe oben erscheinen. In diesem Fall müsste der Anteil  $p$ , mit dem die Zahl 1 erscheint, also  $1/6$  betragen. Ist der Würfel dagegen gezinkt, kann die 1 sowohl häufiger als auch seltener als bei einem Sechstel aller Würfe erscheinen. Die Hypothesen für den Test lauten also  $H_0: p = 1/6$  gegen  $H_1: p \neq 1/6$ .

## Lösung Kapitel 18

Nein, nicht unbedingt. Sie sollten davon ausgehen, dass die Ergebnisse der Befragung von den tatsächlichen Wahlergebnissen abweichen. Entscheidend ist daher die Größe des Stichprobenfehlers. Bei einer Schwankungsbreite von 2% könnte Koalition A zwischen 46% und 50% der Stimmen erhalten, während der Stimmenanteil für Koalition B zwischen 50% und 54% liegen könnte. Bei einer noch größeren Schwankungsbreite würde eine noch stärkere



Überlappung der Wahlergebnisse für die beiden Koalitionen im Bereich des Wahrscheinlichen liegen. Ohne nähere Informationen über die Stichprobengröße oder den Stichprobenfehler können Sie daher nicht erkennen, ob die in der Stichprobe beobachteten Unterschiede zwischen den Stimmenanteilen der beiden Koalitionen auch statistisch signifikant sind und damit einen entsprechenden Rückschluss auf die Grundgesamtheit zulassen (siehe hierzu auch Kapitel 10).

## Lösung Kapitel 19

---

Die Studie basiert auf einer Stichprobe, die ausschließlich junge und gesunde Männer mit einer bewussten und gesunden Ernährung umfasst. Die Schlussfolgerung, die aus den Ergebnissen der Studie gezogen wurde, schränkt die Aussage, ein Ei am Tag habe keinen Einfluss auf den Cholesterinspiegel, jedoch nicht auf diesen Personenkreis der gesunden, jungen Männer mit guter Ernährung ein und impliziert damit, das Ergebnis sei allgemeingültig. Eine derart allgemeine Aussage lässt sich aus einem so beschränkten Experiment jedoch nicht ableiten.

## Lösung Kapitel 20

---

Ein Streudiagramm, das keinen linearen Zusammenhang zeigt, stellt typischerweise eine diffuse Wolke von Datenpunkten dar, die kein klares Muster erkennen lässt. Denkbar sind aber auch andere Verläufe der Datenpunkte; so kann es durchaus sein, dass die Punkte in dem Diagramm ein klares Muster bilden (beispielsweise eine Parabel formen), dieses Muster aber keinen linearen Zusammenhang, sondern zum Beispiel einen quadratischen Zusammenhang darstellt.