

## Meine Hilfe bei statistischen Analysen für wissenschaftliche Arbeiten

Sie schreiben eine wissenschaftliche Arbeit und müssen hierfür Daten  
statistisch analysieren?



Sie haben bisher wenig Erfahrung im Umgang mit der Statistik?



Sie wollen oder können sich nicht selbst in alle Einzelheiten der Statistik einarbeiten?



Sie müssen aus Zeitgründen die statistischen Analysen extern vergeben?



**Dann sind Sie bei Statistik-Service richtig!**



Ich berate und unterstütze Sie bei allen Schritten professionell und fachkundig bzw.  
übernehme die statistische Analyse Ihrer Daten, damit Ihre Arbeit „Hand und  
Fuß“ hat und Sie Ihre Ergebnisse souverän darstellen und präsentieren können.



**Kontakt:**

[info@statistik-service.de](mailto:info@statistik-service.de)

☎ +49 (0)30 443 426 68; +49 (0) 151 6750 7838

Zionskirchstraße 27

**D-10119 Berlin**

## Inhalt

Häufig angewandte statistische Analysemethoden .....	3
Häufig verwendete Grafiken.....	3
Statistische Analyse der aufbereiteten Daten .....	4
Beispielfragestellung 1 .....	4
Beispielfragestellung 2.....	6
Meine angebotenen Leistungen.....	8
Abgrenzung zum „Ghostwriting“ .....	8
Steuerliche Absetzbarkeit des Honorars .....	8
Ablauf.....	8
Preis.....	9
Was können Sie vorab tun? .....	9

## Häufig angewandte statistische Analysemethoden

Die aufgelisteten statistischen Methoden sind lediglich ein Ausschnitt der am häufigsten verwendeten Verfahren. Es gibt auch viele andere, für die ich das Know-how besitze, um Ihre Hypothesen korrekt zu testen.

- Chi-Quadrat-Test
- Korrelationskoeffizient
- t-Test
- Mann-Whitney-U-Test
- ANOVA
- Kruskal-Wallis-Test
- GLM, GZLM, GEE
- Wilcoxon-Test
- Friedman-Test
- Lineare multiple Regression
- Logistische Regression
- Überlebensanalyse, Kaplan-Meier-Methodik, Cox-Regression

## Häufig verwendete Grafiken

- Balkendiagramm
- Liniendiagramm
- Fehlerbalkendiagramm
- Tortendiagramm
- Boxplot
- Histogramm
- Streudiagramm
- Blasendiagramm
- Überlebenskurve
- ROC-Kurve

Auch hier gibt es noch andere Grafiken, welche ich auf Wunsch gerne erzeugen.

Die Farbstile, Schriftgrößen, Schriftarten, usw. können selbstverständlich von Ihnen festgelegt werden. Wenn möglich, liefere ich die Grafiken als Vektorgrafiken, die Sie in der Regel noch ändern können. Eine nachträgliche Änderung von Formaten ist aber auch durch uns jederzeit kostenfrei möglich.

Falls Sie sich nicht sicher sind, welche Grafik für die adäquate Darstellung des Sachverhaltes verwendet werden soll, berate ich Sie gerne.

Für die Analysen verwende ich meistens die Statistik-Software SPSS. Die Grafiken erstelle ich auch mit Excel, PowerPoint, SigmaPlot oder RegioGraph.

## Statistische Analyse der aufbereiteten Daten

Anhand von zwei einfachen statistischen Fragestellungen soll hier kurz dargestellt werden, wie die gelieferte Leistung etwa aussehen könnte.

Die Beispiele basieren auf dem öffentlichen Datensatz „Hyper.sav“ (von Bühl & Zöfel).

### Beispielfragestellung 1

#### Formulierung der Fragestellung

Unterscheidet sich die Medikamentengabe (Alphasan und Betasan) zwischen Männern und Frauen?

Aus der Fragestellung ergibt sich folgende Konstellation: Die unabhängige Größe ist das Geschlecht, die abhängige Größe die Medikamentengabe.

#### Statistische Methoden

Für diesen Fall ist der Chi-Quadrat-Test die richtige statistische Methode zur Prüfung der Unterschiede auf statistische Signifikanz.

Der Chi-Quadrat-Test ist ein statistisches Testverfahren, welches zwei nominale Merkmale auf stochastische Unabhängigkeit untersucht.

Ist der berechnete Signifikanzwert - auch p-Wert genannt - kleiner oder gleich 0,05, dann konnte ein stochastischer Zusammenhang signifikant nachgewiesen werden.

#### Literaturangabe

J. Bortz, G. A. Lienert, Kurzgefasste Statistik für die klinische Forschung – Leitfaden für die verteilungsfreie Analyse kleiner Stichproben, 2.Auflage, Springer 2003, ISBN 3-540-00069-0, S. 82 ff, S. 103 ff.

#### Deskriptive Statistiken

Zur Darstellung sollten hier absolute und relative Häufigkeiten gewählt werden. Mit „N“ wird die Anzahl der gültigen Fälle bezeichnet. Die Spalte „Gesamt“ enthält die Gesamtheit aller Fälle, also ohne Trennung nach Männern und Frauen. Die deskriptiven Statistiken dienen dazu, die Art des Zusammenhangs zu beschreiben.

	Geschlecht					
	Gesamt		männlich		weiblich	
	%	N	%	N	%	N
Alphasan	50%	87	51%	30	50%	57
Betasan	50%	87	49%	29	50%	58
Summe	100%	174	100%	59	100%	115

#### Statistischer Test und Signifikanz

Die Spalte „Signifikanz“ enthält den mit der Statistik-Software berechneten p-Wert und somit das Entscheidungskriterium des Tests.

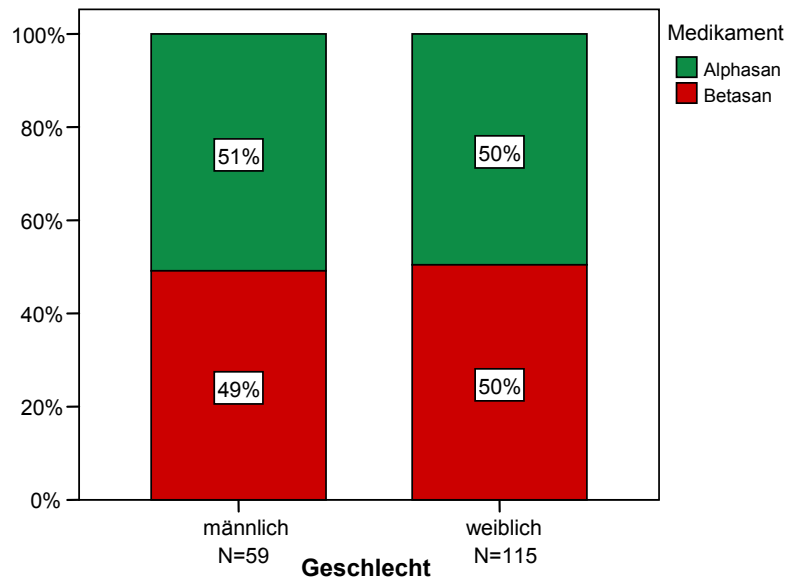
Test	N	Signifikanz
Chi-Quadrat-Test	174	1,000

## Interpretation des Testergebnisses

Ein statistisch signifikanter Unterschied zwischen Männern und Frauen bezüglich der Medikamentengabe ist auf Basis der hier vorliegenden Stichprobe nicht nachweisbar (Chi-Quadrat-Test, p-Wert = 1,00 > 0,05, N = 174). Sowohl bei Männern als auch bei den Frauen liegt eine Gleichverteilung bezüglich der Medikamentengabe vor.

## Grafik

Die Zusammenhänge zwischen den Merkmalen Geschlecht und Medikament werden hier mittels eines Stapelbalkendiagramms dargestellt:



Gut lässt sich hier erkennen, dass sich Männer und Frauen bezüglich der Medikamentengabe nicht unterscheiden. Es wurden 59 Männer und 115 Frauen mit gültiger Medikamentengabe ausgewertet.

## Beispielfragestellung 2

### Formulierung der Fragestellung

Gibt es Gewichtsunterschiede zwischen Männern und Frauen?

Aus der Fragestellung ergibt sich folgende Konstellation: Die unabhängige Größe ist das Geschlecht, die abhängige Größe das Körpergewicht [kg].

### Statistische Methode

Bei vorliegender Normalverteilung in den Gruppen kommt für diesen Fall der t-Test in Frage. Liegt keine Normalverteilung vor, dann sollte der Mann-Whitney-U-Test verwendet werden.

Auf Basis eines statistischen Test und einer grafischen explorativen Analyse muss die Annahme der Normalverteilung abgelehnt werden. Darum wird hier der Mann-Whitney-U-Test verwendet. Der Mann-Whitney-U-Test ist ein nichtparametrisches statistisches Testverfahren, welches für zwei unabhängige Stichproben überprüft, ob sich die zentrale Tendenz der Verteilungen für eine (mindestens) ordinale Zielvariable in den Gruppen unterscheidet.

Ist die berechnete Signifikanz  $p$  kleiner oder gleich 0,05, dann gilt: Ein statistisch signifikant Unterschied in der zentralen Tendenz konnte nachgewiesen werden.

### Literaturangabe

J. Bortz, G. A. Lienert, Kurzgefasste Statistik für die klinische Forschung – Leitfaden für die verteilungsfreie Analyse kleiner Stichproben, Springer 2003, ISBN 3-540-00069-0, S.138 ff.

### Deskriptive Statistiken

Zur Darstellung werden hier Mittelwerte, Standardabweichungen, Median, Minimum und Maximum gewählt. Mit „N“ wird die Anzahl der gültigen Fälle bezeichnet. Die Spalte „Gesamt“ enthält die Gesamtheit aller gültigen Fälle, also ohne Trennung nach Männern und Frauen. Die deskriptiven Statistiken dienen dazu, die Art des Zusammenhangs bzw. Unterschiedes zu beschreiben.

Körpergewicht in kg	Geschlecht		
	Gesamt	männlich	weiblich
Mittelwert	73,8	77,8	71,7
Standardabweichung	12,6	12,1	12,5
Median	72,5	75,0	70,0
Minimum	48	48	48
Maximum	129	124	129
N	174	59	115

Auf Basis der Mittelwerte erkennt man hier, dass in dieser Stichprobe die Männer im Durchschnitt ca. 6 kg schwerer sind als Frauen. Ob dieser Unterschied aber auch statistisch signifikant ist, kann mit Hilfe eines statistischen Tests geprüft werden.

## Statistischer Test und Signifikanz

Die Spalte „Signifikanz“ enthält den mit der Statistiksoftware berechneten p-Wert und somit das Entscheidungskriterium des Tests.

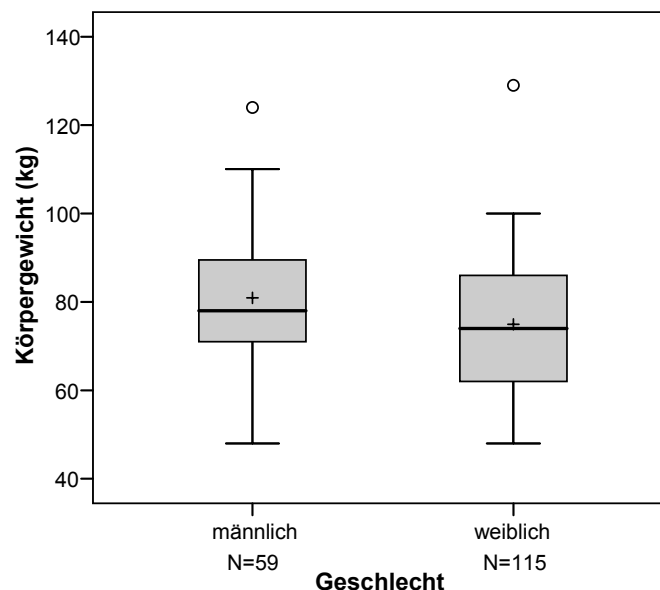
Test	N	Signifikanz
M-W-U-Test	174	0,000

## Interpretation des Testergebnisses

Auf Basis der hier vorliegenden Stichprobe konnte ein statistisch signifikanter Unterschied zwischen Männern und Frauen bezüglich des Körpergewichts zweifelsfrei nachgewiesen werden (M-W-U-Test, p-Wert < 0,001, N = 174). Männer sind im Mittel ca. 78 kg schwer, Frauen ca. 72 kg.

## Grafik

Die Unterschiede im Körpergewicht für Männer und Frauen können hier mittels Boxplots dargestellt werden.



Der hier dargestellte Boxplot enthält sowohl den Median als fetten waagerechten Strich in der Mitte der Box als auch den Mittelwert gekennzeichnet mit einem Plus (+). Der Gewichtsmedian der Männer liegt etwas über dem der Frauen. Gleiches gilt für den Mittelwert. Die genauen Werte können der Tabelle **Deskriptive Statistiken** entnommen werden.

## Meine angebotenen Leistungen

- Beratung zu den statistischen Methoden.
- Beratung zur notwendigen Mindestfallzahl.
- Beratung zu Struktur der notwendigen Datendatei.
- Interpretation der berechneten Testergebnisse.
- Angabe von Literaturstellen zu den verwendeten statistischen Methoden.
- Erläuterung der statistischen Methoden.
- Berechnung der statistischen Tests.
- Tabellarische Darstellung von deskriptiven Statistiken.
- Grafische Veranschaulichung der Ergebnisse.
- Die Ergebnisse können wahlweise in Word, Excel, PowerPoint oder als SPSS-Ausgabe zur Verfügung gestellt werden.
- Lieferung des aufbereiteten Datensatzes in einem vorher festgelegten Format.
- Support zu den verwendeten statistischen Methoden, bzw. berechneten Ergebnissen. Ich lassen Sie auch nach Lieferung der Ergebnisse nicht im Regen stehen.

## Abgrenzung zum „Ghostwriting“

Ich unterstütze Sie professionell bei der statistischen Analyse Ihrer Daten, schreibe aber nicht Ihre wissenschaftliche Arbeit.

Die Eigentumsrechte an den Ergebnissen gehen nach Zahlungseingang komplett an Sie über. Soweit das möglich ist, beraten ich Sie auch bei der inhaltlichen Interpretation der statistischen Ergebnisse. Das geht aber nur im engen Dialog mit Ihnen. Die Basis unserer Zusammenarbeit ist gegenseitiges Vertrauen und Diskretion. Ihre Daten und Ergebnisse werden absolut vertraulich behandelt und nicht an Dritte weitergegeben.

## Steuerliche Absetzbarkeit des Honorars

Die Voraussetzung dafür ist, dass Sie überhaupt Steuern zahlen und jährlich eine Steuererklärung abgeben. Ob hier etwas machbar ist, hängt sehr von Ihrer persönlichen Situation ab und wird wohl auch nicht in allen Bundesländern einheitlich gehandhabt. Am besten, Sie fragen Ihren Steuerberater. In jedem Fall bekommen Sie eine Rechnung mit ausgewiesener Mehrwertsteuer.

## Ablauf

Unsere Zusammenarbeit beginnt in der Regel mit einem ersten Kontaktgespräch. Das kann persönlich in meinem Büro oder per Telefon erfolgen. Dazu sollten Sie Ihre Hypothesen, den auszuwertenden Datensatz bzw. den Fragebogen, und eventuell mögliche Vergleichsarbeiten mitbringen bzw. vorher mailen.

Auf Basis dieses Gesprächs und der von Ihnen zur Verfügung gestellten Unterlagen erstelle ich ein verbindliches Angebot.

Das Angebot enthält den Preis sowie den benötigten Bearbeitungszeitraum. Weiterhin werden die Leistungen so gegliedert, dass Sie auch nur Teile davon in Anspruch nehmen können, wenn dies möglich ist. Z. B. eine Auswertung nur mit den berechneten statistischen Größen und Tests ohne Grafiken, usw.



## Preis

Unsere Leistungen berechnen wir nach unserem zeitlichen Aufwand und dem Schwierigkeitsgrad. Das Gesamthonorar ergibt sich aus dem Angebot. Das aktuelle Stundenhonorar können Sie bei uns gerne telefonisch oder per Email erfragen.

Unser Aufwand wird von folgenden Parametern beeinflusst:

- Form und Qualität der Datendatei. Achten Sie darauf, dass die Daten vollständig sind. Müssen später zusätzliche Daten nachgeliefert werden, dann fallen meist zusätzliche Kosten an. Ihre Daten sollten möglichst keine Eingabefehler enthalten.
- Anzahl und Komplexität Ihrer Hypothesen. Die Anzahl der Fälle hat keine Auswirkung auf das Honorar. Wichtig ist, dass Sie Ihre Hypothesen bzw. Fragestellungen mit Ihrem/r Betreuer/in abgestimmt haben, damit nicht unnötige Kosten anfallen.

## Was können Sie vorab tun?

- Saubere Aufbereitung der zu analysierenden Daten. Damit die berechneten statistischen Größen valide sind, müssen die Daten der Stichprobe ordentlich und sorgfältig aufbereitet werden. Am günstigsten ist es, wenn die Daten schon in einem passenden Format für eine Statistik-Software vorliegen, z. B. als csv-, Excel- oder SPSS-Datendatei.
- Formulierung der zu untersuchenden Fragestellungen, vgl. hierzu Punkt „**Beispielfragestellungen | Formulierung der Fragestellung**“. Die inhaltlichen Fragestellungen (Hypothesen), die auf statistische Signifikanz geprüft werden sollen, sollten von Ihnen möglichst selbst definiert werden. Hier kann ich Sie auch beraten.
- Definition der abhängigen und unabhängigen Größen, vgl. hierzu ebenfalls Punkt „**Beispielfragestellungen | Formulierung der Fragestellung**“. Bei klarer Formulierung der Hypothesen sind abhängige und unabhängige Größen in der Regel erkennbar.
- Sinnvoll ist es natürlich, möglichst viele Beobachtungseinheiten zu erheben, da die Signifikanz von der Fallzahl abhängig ist. Als Faustregel gilt: die kleinste relevante Untersuchungsgruppe sollte etwa 20 Fälle enthalten. Das ist jedoch kein Dogma. Auch für kleinere Fallzahlen hält die schließende Statistik spezielle Methoden bereit. Die minimale Fallzahl für eine Studie kann mit Hilfe von statistischen Methoden bestimmt werden. Auch hierbei kann ich Sie unterstützen.

**Ich freue mich auf Ihren Kontakt!**

**Ihr Dr. Johannes Gladitz**