Bivariate Chi-Quadrat-Verfahren

Inhaltsverzeichnis

Bivariate Chi-Quadrat-Verfahren

Lernhinweise

In diesem Lernschritt soll auf die Bedeutung der

χ^2

-Verfahren im Zusammenhang mit bivariaten Datensätzen eingegangen werden. Neben dem Vergleich eines empirisch erhobenen und eines theoretisch erwarteten bivariaten Datensatzes auf Nominalskalenniveau wird es möglich, zwei nominal skalierte Variablen auf stochastische Unabhängigkeit zu prüfen.

- Grundkonzept aller Verfahren zum Vergleich ganzer Datenverteilungen aus unabhängigen Erhebungen
- Der Vergleich einer empirischen mit einer theoretischen Verteilung
- Vergleich einer empirischen Verteilung mit einer Normalverteilung
- Der Vergleich von zwei empirischen Verteilungen

Benötigte Vorkenntnisse

- Eigenschaften der Skalenniveaus
- Grundprinzip entscheidungsstatistischer Verfahren
- Univariate

  χ^2

  -Verfahren für Erhebungen ohne Messwiederholungen.

Geschätzte Bearbeitungszeit
Nach Konsultation der in Ihrem Curriculum vorgesehenen Vorbereitungslektüre können Sie diesen Lernschritt in ca. 30 Minuten bearbeiten

Hinweise zur Bearbeitung
Den Studierenden im Grundkurs zur Statistik wird empfohlen, den Lernschritt in der vorgegebenen Struktur zu bearbeiten.

Einführung

Die univariaten

χ^2

-Verfahren erlauben uns den Vergleich von zwei univariaten Verteilungen. Dabei werden die auf entsprechende Ausprägungskategorien entfallenden Häufigkeiten verglichen und für die Definition der Prüfgrösse

χ^2

genutzt.

Dieses Verfahren kann für den Vergleich bivariater Verteilungen direkt übernommen werden, wenn wir die Besetzungshäufigkeiten der entsprechenden Ausprägungskombinationen der beiden bivariaten Verteilungen vergleichen.

Theorie (1-3)

Inhaltsübersicht

1. Kontingenztafeln

Interessieren von einer Person, oder allgemeiner gesprochen, von einem System die Ausprägungen von zwei Merkmalen, so sprechen wir von bivariaten Datenerhebungen und bivariaten Datensätzen. Bivariate Datensätze umfassen immer gepaarte Daten, da immer zwei Merkmalsausprägungen von einer Person/einem System stammen und damit zusammengehören.

Probanden Nr.	**Merkmal A: Schuhgrösse**	**Merkmal B: Körperlänge**
1	42	182
2	37	153
3	39	170
etc.	...	...

Wollen wir die in einer Stichprobe erhobenen bivariaten Daten übersichtlich darstellen, so ist dies, unabhängig davon, auf welchem Skalenniveau die Merkmale eingeschätzt wurden, anhand von Tabellen, sogenannten Kontingenztafeln, möglich. Kontingenztafeln werden auch als k×l-Felder-Tafeln bezeichnet wobei k die Anzahl Ausprägungskategorien des Merkmals A und l die Anzahl Ausprägungskategorien des Merkmals B bezeichnet.
Kategorisieren wir die Daten unseres Beispiels, so könnte eine Kontingenztafel wie folgt aussehen:

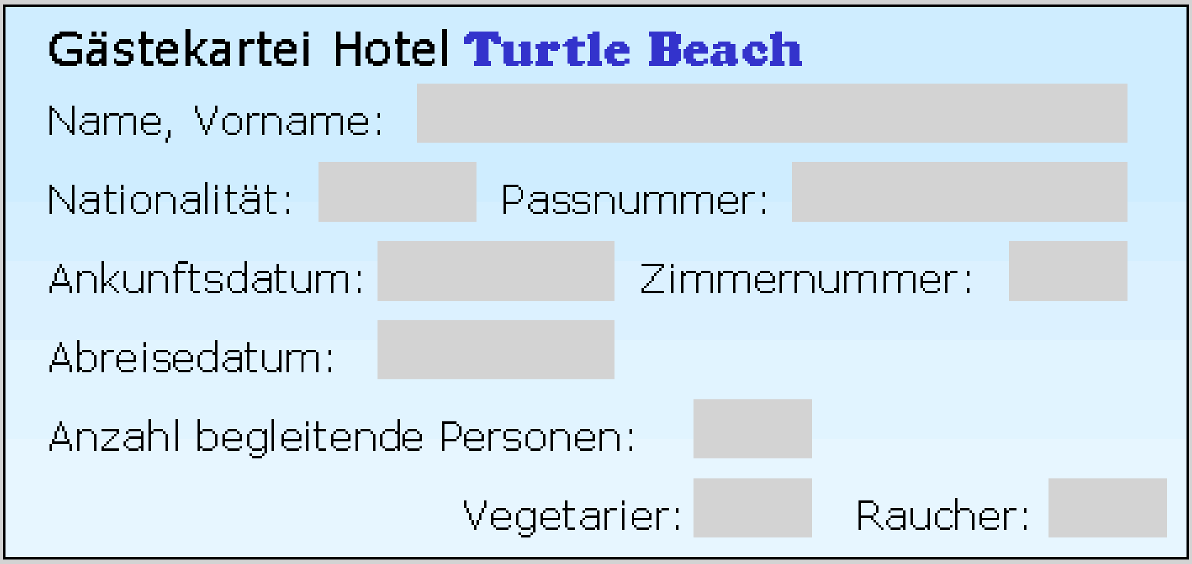
	B: Körpergrösse in cm					
A: Schuhgrösse	**<150**	**151-160**	**161-170**	**171-180**	**181-190**	**>190**
<36	44	52	42	33	62	0
36-38	32	66	76	43	17	0
39-41	11	16	84	53	32	5
41-43	0	2	31	66	59	21
>43	1	19	18	55	42	25

Mit der gewählten Kategorisierung haben wir eine 5×6-Felder-Tafel definiert. In den Feldern dieser Tabelle, man spricht auch von Zellen, stehen die absoluten oder relativen Häufigkeiten, mit denen die einzelnen Kombinationen von Merkmalsausprägungen beobachtet wurden. Unserem Beispiel, das die absoluten Häufigkeiten nennt, entnehmen wir, dass die Kombination "Schuhgrösse 39-41" und "Körpergrösse 171-180" für 53 Personen unserer Stichprobe zutrifft.
Wir merken uns:

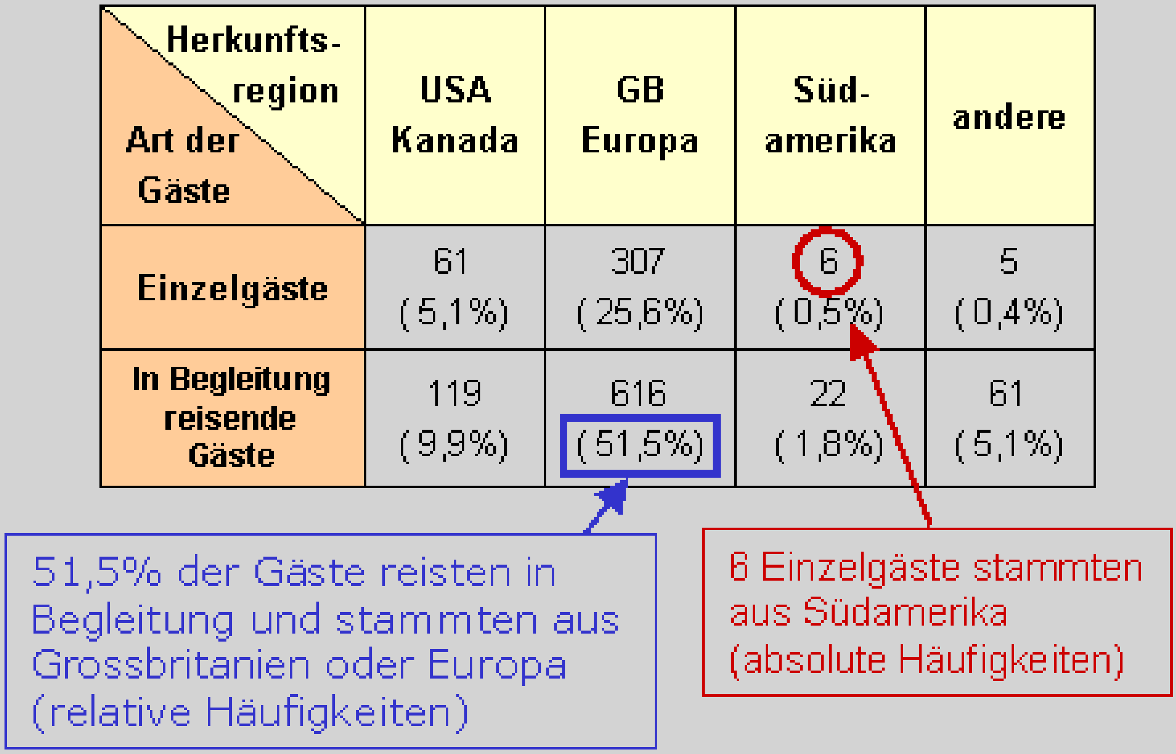
Bivariate Verteilungen können - unabhängig vom Skalenniveau der beiden Merkmale und unabhängig von deren Kategorisierung - anhand von Kontingenztafeln (k×l- Felder-Tafeln) dargestellt werden. Kontingenztafeln werden mit den in diesem Lernschritt zu besprechenden Verfahren bearbeitet, wobei immer auf nominalem Skalenniveau ausgewertet wird.

Beispiel einer Kontingenztafel für ein dichotom und ein nominal skaliertes Merkmal

Beispiel einer Kontingenztafel für ein dichotom und ein nominal skaliertes Merkmal
Der Gästedatei des Hotels Turtle Beach umfasst die in der Hotellerie üblichen Angaben



Vizedirektor Jürgen Schmählich plant die Werbekampagne für die nächste Saison. In diesem Zusammenhang erstellt er aus den Daten des letzten Geschäftsjahres eine Kontingenztafel mit den Merkmalen "Art der Gäste" und "Herkunftsregion der Gäste".



Art der Gäste \ Herkunfts-region	USA Kanada	GB Europa	Süd-amerika	andere
Einzelgäste	61 (5,1%)	307 (25,6%)	6 (0,5%)	5 (0,4%)
In Begleitung reisende Gäste	119 (9,9%)	616 (51,5%)	22 (1,8%)	61 (5,1%)

Die Ausprägungskategorien des Merkmals "Art der Gäste" basiert auf der proportional skalierten Angabe über die Anzahl begleitender Personen. In der Auszählung wurde nur zwischen "Einzelgästen" und "in Begleitung reisende Gäste" unterschieden. Das Merkmal "Art der Gäste" ist damit dichotom skaliert. Das Merkmal "Herkunftsregion" ist nominal skaliert, wobei sich Jürgen Schmählich bei der Festlegung der Kategorien an den wichtigsten Gästegruppen orientierte.

2. Vergleich einer bivariaten mit einer theoretisch erwarteten Verteilung

Liegt eine empirisch ermittelte bivariate Verteilung in Form einer Kontingenztafel vor, so kann diese in einfacher Weise mit einer theoretisch erwarteten bivariaten Verteilung verglichen werden. Wie das folgende **Beispiel** zeigt, stützen wir uns dabei auf das klassische

χ^2

-Verfahren.

Unser Interesse gilt der Verteilung der beruflichen Positionen auf die Mitarbeiterinnen und Mitarbeiter in der Euro-Test AG. Für die berufliche Position wurden nur drei Kategorien gewählt so dass die Daten in einer (k×l = 2×3) Kontingenztafel darstellbar sind. In die Tafelzellen eingetragen sind die beobachteten absoluten Häufigkeiten f_{bij}.

	Oberes Kader	Mittleres Kader	Angestellte	Σ
Frauen	8 (f_{b11})	25 (f_{b12})	14 (f_{b13})	47
Männer	7 (f_{b21})	18 (f_{b22})	8 (f_{b23})	33
Σ	**15**	**43**	**22**	80

Vor dem Hintergrund der Forderung nach Chancengleichheit der Geschlechter stellt sich die Frage, inwieweit Männer und Frauen in den Kategorien "Oberes Kader", "Mittleres Kader" und "Angestellte" gleich stark vertreten sind.

Zur Untersuchung dieser Frage postulieren wir anhand der Kolonnensummen die bei absoluter Chancengleichheit theoretisch zu erwartende Verteilung: Bei idealer Gleichstellung von Männern und Frauen sind diese in den einzelnen Positions-Kategorien gleichverteilt. So erhalten wir die folgende **theoretisch erwartete Verteilung** der absoluten Häufigkeiten f_{eij}:

	Oberes Kader	Mittleres Kader	Angestellte	Σ
Frauen	**7,5** (f_{e11})	**21,5** (f_{e12})	**11** (f_{e13})	40
Männer	**7,5** (f_{e21})	**21,5** (f_{e22})	**11** (f_{e23})	40
Σ	**15**	**43**	**22**	80

Die empirische und die theoretische Verteilung können wir nun mit dem klassischen

χ^2

-Verfahren vergleichen. Dabei schätzen wir ab, ob die Unterschiede zwischen den beiden Verteilungen so gravierend sind, dass wir bei der Interpretation der Unterschiede den Zufall mit einer Irrtumswahrscheinlichkeit $p < 5\%$ ausschliessen können.

Überlegen Sie sich die für diesen Vergleich notwendigen Arbeitsschritte, bevor Sie zur nächsten Seite übergehen.

Wir betrachten die Arbeitsschritte im Detail:

1. **Voraussetzungen und Wahl des Prüfverfahrens:** Wir möchten eine empirisch erhobene mit einer theoretisch zu erwartenden bivariaten Verteilung vergleichen. Das eine Merkmal ist nominal (konkret dichotom), das andere Merkmal ordinal skaliert. Da keine der theoretisch zu erwartenden absoluten Häufigkeiten kleiner ist als 5, ist ein klassisches

   χ^2

   -Verfahren angezeigt.
2. **Arbeits- und Alternativhypothese:**
   Arbeitshypothese H_0 : Die empirisch erhobene und die theoretisch zu erwartende Verteilung unterscheiden sich nur zufällig.
   Alternativhypothese H_1 : Die empirische und die theoretisch zu erwartende Verteilung unterscheiden sich nicht zufällig.
3. **Bestimmung der Prüfgrösse**: Den Ausprägungsgrad der Prüfgrösse

   χ^2

   bestimmen wir über den paarweisen Vergleich der absoluten Häufigkeiten in den 6 Zellen (2×3) der empirischen und der theoretischen Verteilung.
   Allgemein bestimmt sich die Prüfgrösse wie folgt:

   $$\chi^2 = \sum_{i=1}^{k} \sum_{j=1}^{l} \frac{(f_{b_{ij}} - f_{e_{ij}})^2}{f_{e_{ij}}}$$

   In unserem konkreten Beispiel ist k = 2 und l = 3. Damit erhalten wir für den Ausprägungsgrad unserer Prüfgrösse:

   $$\chi^2 = \sum_{i=1}^{2} \sum_{j=1}^{3} \frac{(f_{b_{ij}} - f_{e_{ij}})^2}{f_{e_{ij}}} = 2,84$$
4. **Prüfverteilung und Überschreitungswahrscheinlichkeit**: Die Prüfgrösse ist

   χ^2

   -verteilt, wir müssen über den Freiheitsgrad noch festlegen, welche -Verteilung Prüfverteilung ist.
   Wir erinnern uns: Für die Bestimmung des Freiheitsgrades subtrahieren wir von der Anzahl verglichener Ausprägungskategorien (hier k×l) die Anzahl Vorgaben für die theoretische Verteilung.
   In unserem Beispiel haben wir für die Ermittlung der theoretischen Verteilung die Kolonnensummen vorgegeben, d.h. wir haben 3 Vorgaben gemacht:
   df = (k×l) - Anzahl Vorgaben = 2×3 - 3 = 3
   Prüfverteilung ist damit die -Verteilung mit df = 3.
   Der

   χ^2

- Tabelle entnehmen wir, dass bei einem Freiheitsgrad von 3 der -Wert von 2,84 mit einer Wahrscheinlichkeit von mehr als 20% zufällig erreicht oder überschritten wird. Bei dieser Überschreitungswahrscheinlichkeit kann unsere Arbeitshypothese H0 nicht abgelehnt werden.

Chi-Quadrat - Tabelle

5. **Interpretation**: Dieses Resultat bedeutet, dass wir bei der Interpretation der Unterschiede zwischen der theoretischen und der empirischen Verteilung den Zufall nicht ausschliessen können. Mit dieser Untersuchung kann eine Benachteilig des einen oder anderen Geschlechtes in der Euro-Test AG nicht nachgewiesen werden.

3. Spezialfall

Anhand eines

χ^2

-Verfahrens können wir zwei nominal skalierte Variablen auch in einfacher Weise auf stochastische Unabhängigkeit prüfen. Das Prinzip dieses Verfahrens lässt sich wie folgt zusammenfassen:

1. **Ausgangspunkt** ist eine empirisch erhobene Kontingenztafel (k×l-Feldertafel), in der die absoluten Häufigkeiten der beobachteten Ausprägungskombinationen von zwei Variablen zusammengetragen sind.
2. Anhand einfacher **wahrscheinlichkeitstheoretischer Überlegungen** lässt sich daraus eine theoretisch zu erwartende Kontingenztafel für den Fall ermitteln, dass die beiden Variablen stochastisch unabhängig sind.
3. Wir prüfen die empirisch erhobene und die bei stochastischer Unabhängigkeit zu erwartende Kontingenztafel mit einem

χ^2

-Verfahren auf einen signifikanten Unterschied.

Unterscheidet sich empirische und theoretische Kontingenztafel signifikant, so darf mit der entsprechenden Irrtumswahrscheinlichkeit zwischen den beiden Variablen eine stochastische Abhängigkeit angenommen werden.

Unterscheiden sich die beiden Kontingenztafel nicht signifikant, so kann nicht ausgeschlossen werden, dass die beiden Variablen stochastisch unabhängig sind.

In der Regel setzen wir für derartige Datenanalysen ein Statistikprogramm ein (z.B. SPSS). Das grundsätzliche Vorgehen soll aber am einfachsten Fall, einer 4-Feldertafel, illustriert werden. Klicken Sie im obenstehenden Diagramm auf die Sie interessierenden Aspekte.

Analyse einer Kontingenztafel

Kann zwischen den zwei Variablen einer

$k \times l$

-Kontingenztafel eine stochastische Abhängigkeit nachgewiesen werden, so stellt sich die Frage nach der inhaltlichen Interpretation dieses Resultates; in der Regel möchten wir den Zusammenhang zwischen den Variablen inhaltlich erläutern. Im folgenden werden zwei Wege aufgezeigt, die die inhaltliche Interpretation eines signifikanten stochastischen Zusammenhangs unterstützen.

Vergleich der Kolonnen- oder der Zeilenprozentsätze mit den entsprechenden Randprozentsätzen (Manuelle Berechnungen möglich)

Das folgende einfache Beispiel zeigt, was wir unter Kolonnen-Randsummen, Zeilen-Randsummen, Kolonnen-Prozentwerten und Zeilen-Randsummen-Prozentwerten verstehen.

Dieses Element (Animation, Video etc.) kann in der PDF version nicht dargestellt werden und ist nur in der online Version sichtbar. [link]

Der Vergleich der Kolonnen-Prozentwerte (Prozentwerte in den Zellen der Kontingenztafel) mit den zugehörigen Zeilen-Randsummen-Prozentwerten kann uns die Interpretation des nachgewiesenen stochastischen Zusammenhangs erleichtern. Dazu folgende Überlegungen:

Bei stochastischer Unabhängigkeit der Variablen A und B stimmen die Kolonnen-Prozentwerte der einzelnen Zellen mit den entsprechenden Zeilen-Randsummen-Prozentwerten exakt überein. Bei stochastischer Abhängigkeit der Variablen A und B hingegen zeigen sich markante Differenzen; auf diese ist bei der inhaltlichen Interpretation insbesondere zu achten.

Zwei extreme Beispiele sollen dies veranschaulichen. Wählen Sie erst das Beispiel und dann die Arbeitsschritte.

Dieses Element (Animation, Video etc.) kann in der PDF version nicht dargestellt werden und ist nur in der online Version sichtbar. [link]

In Fallbeispiel 3 wird dieses Verfahren für die inhaltliche Interpretation einer signifikanten Kontingenztafel eingesetzt.
Vergleich der Standard-Residualwerte der einzelnen Zellen der Kontingenztafel (Berechnungen in der Regel mit einem Statistikprogramm, z.B. SPSS)
Der sog. Standard-Residualwert [Standardized Residual] wird für jede Zelle der Kontingenztafel bestimmt. Er beschreibt in standardisierter Form das Mass der Übereinstimmung zwischen der tatsächlich beobachteten Besetzungshäufigkeit und der bei stochastischer Unabhängigkeit der beiden Variablen theoretisch zu erwartenden Besetzungshäufigkeit.
Standard-Residualwert für die Zelle (a_i, b_j) :

$$sRes_{ij} = \frac{f_{b_{ij}} - f_{e_{ij}}}{\sqrt{f_{e_{ij}}}}$$

Bei grossen Stichproben, das glauben wir den Statistikern gerne, entspricht sResij in etwa einem z-Wert. Dies erlaubt die Einschätzung der Zufälligkeit eines Ausprägungsgrades von sResij. Auf der geschätzten Überschreitungswahrscheinlichkeit basiert die folgende Faustregel

- $sRes_{ij} > 1{,}96$ oder $sRes_{ij} < -1{,}96$:
  Die Abweichung zwischen der beobachteten und der erwarteten Häufigkeit in der Zelle (a_i, b_j) ist auffällig ("signifikant"), die Zelle ist für die Interpretation des stochastischen Zusammenhangs bedeutungsvoll.
- $sRes_{ij} < 1{,}96$ oder $sRes_{ij} > -1{,}96$:
  Die Differenz zwischen beobachteter und erwarteter Häufigkeit ist nicht auffällig. Die Zelle (a_i, b_j) trägt zum stochastischen Zusammenhang bei.

Dieses Element (Animation, Video etc.) kann in der PDF version nicht dargestellt werden und ist nur in der online Version sichtbar. [link]

Da in die Berechnung der Standard-Residualwerte die bei stochastischer Unabhängigkeit theoretisch erwarteten Häufigkeiten

$$f_{e_{ij}}$$

einfliessen, diese manuell aber nur umständlich zu bestimmen sind, lassen wir uns die Standard-Residualwerte von einem Statistikprogramm (z.B. SPSS) ausgeben.
Im Fallbeispiel 3 wird ein signifikanter stochastischer Zusammenhang auch anhand der Standard-Residualwerte interpretiert resp. erläutert.

Fallbeispiel (1-3)

- 1. Stochastische Unabhängigkeit dichotom skalierter Merkmale
- 2. Stochastische Unabhängigkeit mehrfach gestufter Merkmale
- 3. Stochastische Unabhängigkeit und Analyse der Kontingenztafel (mehrfach gestufter Merkmale)

1. Stochastische Unabhängigkeit dichotom skalierter Merkmale

Besteht zwischen 'Geschlech' und 'erstbestelltem Drink' ein stochastischer Zusammenhang?
Anderson C. Flash ist Barkeeper im Turtle-Beach Hotel (Tobago, Karibik) und hat alle Hände voll zu tun. Jürgen Schmählich, der Hoteldirektor, hat die neu angekommenen Gäste zu einem Apéro eingeladen, an dem Andersons beide berühmtesten Drinks ("ACF-Punch" und "Taifun") ausgeschenkt werden.
Als aufmerksamer und geschäftstüchtiger Barkeeper möchte Anderson herausfinden, ob die Namen seiner beiden Kreationen die Damen und Herren in vergleichbarer Weise ansprechen, oder ob geschlechtsspezifische Präferenzen bestehen.
Zu diesem Zweck notiert er sich bei jedem Gast das Geschlecht und welcher der beiden Drinks als erster bestellt wird. Nach Dienstschluss trägt Anderson diese Daten in einer Kontingenztafel (konkret einer 4-Feldertafel) zusammen.

Erstbestellung

Geschlecht

	ACF-Punch	Taifun	Σ
Herren	21	8	29
Damen	4	11	15
Σ	25	19	

Da Anderson auch auf seiner Ferieninsel Zugriff auf einen web-gestützten Statistikkurs hat, fällt ihm die Bearbeitung dieser Daten leicht.

1. **Fragestellung**:
   Stehen die beiden Variablen "Geschlecht" und "erstbestellter Drink" in einem stochastischen Zusammenhang?
2. **Datenstruktur und Wahl des Prüfverfahrens**:
   Beide Variablen sind dichotom skaliert und wurden in einer Stichprobe von 44 Gästen erhoben. Bei der gegebenen Grösse der ermittelten Häufigkeiten in den Zellen der 4-Feldertafel darf angenommen werden, dass die bei stochastischer Unabhängigkeit zu erwartenden theoretischen Häufigkeiten alle grösser sind als 5. Angezeigt ist damit ein chi-quadrat.gif' , 'Height= -Verfahren zur Prüfung der stochastischen Unabhängigkeit von zwei dichotom skalierten Variablen.
3. **Arbeitshypothese H_0** :
   Die beiden Variablen sind stochastisch unabhängig.
   Alternativhypothese H_1 :

Die beiden Variablen sind stochastisch abhängig.

4. **Bestimmung der Prüfgrösse**:

$$\chi^2 = \frac{n*(a*d-b*c)^2}{(a+b)*(c+d)*(a+c)*(b+d)} = \frac{44*(21*11-8*4)^2}{29*15*25*19} = 8{,}43$$

5. **Prüfverteilung und Überschreitungswahrscheinlichkeit**:
Prüfverteilung ist die

$$\chi^2$$

-Verteilung mit dem Freiheitsgrad df=1. Der entsprechenden

$$\chi^2$$

-Tabelle entnehmen wir, dass der ermittelte Ausprägungsgrad nur noch in weniger als 1% der Fälle zufällig erreicht oder überschritten wird. Unter diesen Umständen lehnt Anderson seine Arbeitshypothese H_0 mit einer Irrtumswahrscheinlichkeit $p < 1\%$ ab.

6. **Interpretation**:
Mit einer Irrtumswahrscheinlichkeit $p < 1\%$ darf angenommen werden, dass für die erste Wahl des offerierten Drinks eine geschlechtsspezifische Präferenz besteht. Der Kontingenztafel entnimmt Anderson, dass die Herren offensichtlich den "ACF-Punch" bevorzugen, während die Damen den "Taifun" vorziehen.

Berechnung mit SPSS zeigen

Aus Neugierde werten wir Andersons Daten noch mit SPSS aus.
Als erstes geben wir seine Daten (Ausprägungen der beiden Variablen "Geschlecht" und "Getränk" für alle 44 Gäste) ein und speichern sie in einem Datenfile mit dem Namen sex-apero.sav.
Mit den folgenden Befehlen lesen wir unser Datenfile ein und werten es im Rahmen einer 4-Feldertafel aus:

```
GET FILE "X:\\SPSS-dat\\sex-apero.sav".
CROSSTABS TABLES = sex BY getraenk /STATISTICS = CHISQ.
```

Geschlecht * Getränk

Count

		Getränk		Total
		ACF-Punch	Taifun	
Geschlecht	Herren	21	8	29
	Damen	4	11	15
Total		25	19	44

Chi-Square Tests

	Value	df	Asymp. Sig. (2-sided)	Exact Sig. (2-sided)	Exact Sig. (1-sided)
Pearson Chi-Square	8.433[b]	1	.004		
Continuity Correction[a]	6.671	1	.010		
Likelihood Ratio	8.617	1	.003		
Fisher's Exact Test				.009	.005
N of Valid Cases	44				

a. Computed only for a 2x2 table

b. 0 cells (.0%) have expected count less than 5. The minimum expected count is 6.48.

Nach der Ausgabe der 4-Feldertafel folgt eine Tabelle mit den Ausprägungen verschiedener Prüfgrössen und deren Überschreitungswahrscheinlichkeiten.

- Die Ausprägung des klassischen

  χ^2

  findet sich unter "Pearson Chi-Square"; sie entspricht dem manuell bestimmten Ausprägungsgrad. Die zweiseitige Überschreitungswahrscheinlichkeit beträgt 0,4%. Auch dieses Resultat entspricht unserer Handauswertung.
- Bei Berücksichtigung der Kontinuitätskorrektur nach Yates (Continuity Correction) fällt die Einschätzung der Überschreitungswahrscheinlichkeit etwas grösser aus, das Resultat bleibt aber zumindest auf dem 5%-Niveau signifikant.
- Unter "Likelihood Ratio" wird eine alternative Prüfgrösse ausgegeben, die sich mit grösser werdender Stichprobe dem

  χ^2

  -Wert annähert.

- Der Fussnote b. entnehmen wir, dass die Voraussetzung des 4-Felder-

  χ^2

  -Verfahrens erfüllt ist und keine der bei stochastischer Unabhängigkeit theoretisch erwarteten absoluten Häufigkeiten kleiner ist als 5.
- Wäre diese Voraussetzung verletzt, so könnten wir noch den Fisher-Yates-Test einsetzen. Die zugehörige Überschreitungswahrscheinlichkeit findet sich unter "Fisher's Exact Test".

Bei der Interpretation des Resultates kommen wir zum selben Schluss wie Anderson: Mit einer Irrtumswahrscheinlichkeit $p < 5\%$ darf angenommen werden, dass für die erste Wahl des offerierten Getränkes eine geschlechtsspezifische Präferenz besteht.

2. Stochastische Unabhängigkeit mehrfach gestufter Merkmale

Besteht zwischen 'Tätigkeitsbereich' und 'Informationsmedium' ein stochastischer Zusammenhang?
Im Rahmen einer Erhebung zu den Lebensgewohnheiten von Lehrlingen wurden in einer Stichprobe von n = 56 Berufsschülern u.a. auch die Merkmale **"Tätigkeitsbereich"** und **"Primär benutztes Informationsmedium"** erhoben. Dabei waren nur je drei nominal skalierte Ausprägungen vorgesehen. Zum Merkmal "Tätigkeitsbereich" konnte zwischen den Ausprägungen "Produktion", "Verwaltung" oder "Dienstleistung" gewählt werden, zum Merkmal "Primär benutztes Informationsmittel" hatten sich die Lehrlinge entweder für "Tagespresse", für "Rundfunk" oder für "Fernsehen" zu entscheiden.
Wir möchten nun prüfen, inwieweit das Merkmal "Tätigkeitsfeld" mit dem Merkmal "Primär benutztes Informationsmittel" in einem stochastischen Zusammenhang steht. Zu diesem Zweck lassen wir die erhobenen Daten in einer Kontingenztafel (konkret in einer 3×3-Feldertafel) zusammentragen und prüfen dann mit einem k×l-

χ^2

-Verfahren die beiden Merkmale auf stochastische Unabhängigkeit.
Wir lösen diese Aufgaben direkt mit SPSS über die folgenden Arbeitsschritte:

- **Datenstruktur und Wahl des Prüfverfahrens**:
  Beide interessierenden Merkmale sind nominal skaliert. Zur Prüfung auf stochastische Abhängigkeit kann ein k×l- chi-quadrat.gif' , 'Height= -Verfahren dienlich sein
- **Arbeitshypothese H_0** : Die beiden Merkmale (Variablen) sind stochastisch unabhängig.

  Alternativhypothese H_1 : Die beiden Merkmale (Variablen) sind stochastisch abhängig.
- **Bestimmung des Ausprägungsgrades der Prüfgrösse und der zugehörigen Überschreitungswahrscheinlichkeit mit SPSS**:
  Wir codieren die Daten, geben sie in SPSS ein und speichern sie in einem Datenfile mit dem Namen info.sav (hier nicht wiedergegeben). Mit den folgenden Befehlen lesen wir das Datenfile ein und verlangen eine Auswertung nach dem k×l-

  χ^2

  Verfahren:

Tätigkeitsbereich * primäres Informationsmedium

Count

		primäres Informationsmedium			Total
		Presse	Rundfunk	Fernsehen	
Tätigkeitsbereich	Produktion	6	3	12	21
	Verwaltung	5	8	5	18
	Dienstleistung	7	5	5	17
Total		18	16	22	56

Chi-Square Tests

	Value	df	Asymp. Sig. (2-sided)
Pearson Chi-Square	6.446[a]	4	.168
Likelihood Ratio	6.427	4	.169
N of Valid Cases	56		

a. 1 cells (11.1%) have expected count less than 5. The minimum expected count is 4.86.

Die Computerausgabe zeigt erst die 3×3-Feldertafel, in der die erhobenen Daten zusammengetragen sind. Dann folgt der Ausprägungsgrad der Prüfgrösse (Pearson Chi-Square) mit der zugehörigen Überschreitungswahrscheinlichkeit.

Unter "Likelihood Ratio" wird eine alternative Prüfgrösse ausgegeben, die sich mit grösser werdender Stichprobe dem

χ^2

-Wert annähert.

In der Fussnote a. ist die Zahl der Ausprägungskombinationen genannt, für die die bei Gültigkeit von H_0 theoretisch zu erwartenden absoluten Häufigkeiten geringer sind als 5. Das Resultat eines k×l-

χ^2

-Verfahrens ist interpretierbar, wenn nicht mehr als 20% der theoretisch erwarteten Häufigkeiten kleiner sind als 5. Im gegebenen Fall trifft dies nur für eine von neun Ausprägungskombinationen zu (11,1%); das Auswertungsresultat kann somit interpretiert werden.

- **Interpretation**:

Die Überschreitungswahrscheinlichkeit der Prüfgrösse beträgt 16,8%, die Arbeitshypothese kann beibehalten werden. Inhaltlich bedeutet dies, dass nicht ausgeschlossen werden kann, dass die beiden interessierenden Merkmale "Tätigkeitsfeld" und "primäres Informationsmittel" stochastisch unabhängig sind. Eine Irrtumswahrscheinlichkeit kann für dieses Resultat nicht angegeben werden.

3. Stochastische Unabhängigkeit und Analyse der Kontingenztafel (mehrfach gestufter Merkmale)

Überprüfung des Erfolges verschiedener Therapieformen

Im Rahmen einer klinischen Wirksamkeitsstudie werden Managerinnen und Manager (Alter zwischen 40 und 50 Jahren, vergleichbare Arbeits- und Lebensumstände) mit demselben psychosomatischen Krankheitsbild nach drei verschiedenen Methoden therapiert. Zur Anwendung kommen folgende Therapieformen: 'Beratungsgespräch', 'Verhaltenstraining kombiniert mit medikamentöser Behandlung' und 'Medikamentöse Behandlung'.

Nach jeweils sechs Monaten wird der Behandlungserfolg in den Kategorien 'gut', 'mittel' und 'gering' eingeschätzt und zusammen mit der Behandlungsmethode im SPSS-Datenfile **methode-erfolg.sav** registriert. Heute liegen von 146 Personen diese Daten vor; an ihnen wollen wir die folgenden Fragen beantworten:

1. Besteht zwischen der ‚Behandlungsmethode' und dem ‚Behandlungserfolg' ein statistisch nachweisbarer Zusammenhang?
2. Wenn ein Zusammenhang nachweisbar ist, stellt sich die Frage, wie dieser bezüglich der Behandlungsmethoden zu interpretieren ist.

Wir beantworten diese Fragen mit Hilfe von SPSS über die folgenden Arbeitsschritte:

1. **Datenstruktur und Wahl des Prüfverfahrens**:
   Die Merkmale 'Behandlungsmethode' und 'Behandlungserfolg' sind nominal skaliert. Zur Prüfung auf stochastische Abhängigkeit wählen wir ein
   $k \times l$
   -
   χ^2
   -Verfahren.
2. **Arbeitshypothese H_0** : Die beiden Merkmale sind stochastisch unabhängig. **Alternativhypothese H_1** : Die beiden Merkmale sind stochastisch abhängig.
3. Ausgabe einer Kontingenztafel mit Kolonnen-Prozentwerten und Standardisierten Residualwerten, Bestimmung des Ausprägungsgrades der Prüfgrösse und der zugehörigen Überschreitungswahrscheinlichkeit (mit SPSS).
   Die Rohdaten sind im Datenfile **methode-erfolg.sav** gespeichert. Wir analysieren sie mit dem folgenden Befehlssatz:
   GET FILE 'X:\\SPSS_DAT**methode-erfolg.sav**'.
   CROSSTABS TABLES = **erfolg** BY **methode** / CELLS = COUNT COLUMN
   SRESID/ STATISTICS = CHISQ.
   Zur Beachtung:

Mit der Spezifikation 'CELLS = COUNT COLUMN SRESID' verlangen wir - im Gegensatz zu den Auswertungen in den vorangegangenen Fallbeispielen - auch die Ausgabe der Kolonnen-Prozentwerte und der Standardisierten Residualwerte.
Die Computerausgabe zeigt uns erst die Kontingenztafel. Dabei sind in jeder Zelle von oben nach unten die folgenden Grössen eingetragen: Absolute Häufigkeit, Kolonnen-Prozentwert und Standard-Residualwert.

SPSS-Output

Erfolg * Methode Crosstabulation

			Methode: Beratung	Methode: Verhaltenstraining -Medikament	Methode: Medikament	Total
Erfolg	gut	Count	10	26	7	43
		% within Methode	21.7%	46.4%	15.9%	29.5%
		Std. Residual	-1.0	2.3	-1.7	
	mittel	Count	16	23	30	69
		% within Methode	34.8%	41.1%	68.2%	47.3%
		Std. Residual	-1.2	-.7	2.0	
	gering	Count	20	7	7	34
		% within Methode	43.5%	12.5%	15.9%	23.3%
		Std. Residual	2.8	-1.7	-1.0	
Total		Count	46	56	44	146
		% within Methode	100.0%	100.0%	100.0%	100.0%

Chi-Square Tests

	Value	df	Asymp. Sig. (2-sided)
Pearson Chi-Square	27.073[a]	4	.000
Likelihood Ratio	25.625	4	.000
Linear-by-Linear Association	2.160	1	.142
N of Valid Cases	146		

a. 0 cells (.0%) have expected count less than 5. The minimum expected count is 10.25.

Im zweiten Teil folgt der Ausprägungsgrad der Prüfgrösse (Pearson Chi-Square) mit der zugehörigen Überschreitungswahrscheinlichkeit. Die Fussnote a zeigt, dass keine der bei Gültigkeit von H_0 theoretisch zu erwartenden absoluten Häufigkeiten kleiner ist als 5; damit kann der Computer-Output interpretiert werden.

4. **Interpretation:**
Zur ersten Frage nach einem stochastischen Zusammenhang:

Die Überschreitungswahrscheinlichkeit für den Ausprägungsgrad der Prüfgrösse ist kleiner als 0,1%, die Arbeitshypothese H_0 kann damit zu Gunsten der Alternativhypothese H_1 abgelehnt werden. Damit steht fest, dass zwischen den Merkmalen 'Behandlungsmethode' und 'Behandlungserfolg' ein hoch signifikanter stochastischer Zusammenhang besteht.

Zur zweiten Frage nach der Charakteristik des Zusammenhangs:

Wir vergleichen die Kolonnen-Prozentwerte mit den zugehörigen Zeilen-Randsummen-Prozentwerten und stellen fest, dass die Kolonnen-Prozentwerte der Ausprägungskombinationen (Verhaltenstraining-Medikament, gut), (Medikament, mittel) und (Beratung, gering) deutlich über den Zeilen-Randsummen-Prozentwerten liegen, alle anderen Kolonnen-Prozentwerte liegen zum Teil markant darunter.

Die Ausprägungskombinationen mit den grössten Abweichungen geben Hinweise, wie der stochastische Zusammenhang inhaltlich zu beurteilen ist.

Zu demselben Resultat führt der Vergleich der Standardisierten Residualwerte. Grosse positive Werte zeigen sich auch hier für die Kombinationen (Verhaltenstraining-Medikament, gut), (Medikament, mittel) und (Beratung, gering). Dies bedeutet, dass diese Zellen im Vergleich zu den bei Gültigkeit der Arbeitshypothese H_0 theoretisch zu erwartenden Belegungshäufigkeiten namhaft überbelegt sind.

Inhaltlich können wir auf Grund dieser Daten den stochastischen Zusammenhang zwischen den Merkmalen 'Behandlungsmethode' und 'Behandlungserfolg' wie folgt charakterisieren: Die Methode 'Verhaltenstraining kombiniert mit einer medikamentösen Behandlung' zeitigt den besten Erfolg, die Methode 'Medikamentöse Behandlung' mehrheitlich einen mittleren und die Methode 'Beratungsgespräch' mehrheitlich einen geringen Erfolg.