Aufgabe 32 a), b):

Bei 32 a+b sollte zumindest diskutiert werden, inwieweit die Punkthypothese überhaupt Sinn macht, also ob es der Supermarktkette darauf ankommt, genau 1500 DM als Durchschnitteinkommen in dieser Käuferschicht zu haben, oder ob bspw. *mindestens* 1500 DM (dann wäre die neg. Gegenhypothese zu testen) nicht besser wären. Bleibt man bei der (positiv formulierten) Punkthypothese, sollte deutlich gemacht werden, daß der Annahmebereich sich mit vergrößerten Signifikanzniveau (also kleinerem α) ebenfalls vergrößert-> erhöhter β-Fehler!

Hypothesentest für Mittelwerte und Standardabweichung

a) geg.: N = 10000 Kunden n = 10 $\bar{x} = 1600 DM$

$\hat{s} = 250 DM$ α = 0,1 bzw. 0,05

- σ ist unbekannt !

- $n \leq 30$ $\Rightarrow \dfrac{\bar{x} - m}{\frac{\hat{s}}{\sqrt{n}}}$ ist t-verteilt

- kein Endlichkeitsfaktor, weil $\dfrac{n}{N} = \dfrac{10}{10000} = 0,001 < 0,05$

Die Hypothese könnte z.B. $H_0\!: m = m_0 = 1500 DM$ (beidseitiger Test) lauten. (Diskussion! s.o.)

Bestimmung der Grenzen des Annahmebereiches:

$$\bar{x}_{r_1,r_2} = m_0 \pm t_{\frac{a_0}{2}} \cdot \frac{\hat{s}}{\sqrt{n}}$$

$$a_0 = 0{,}10 \Rightarrow t_{(\Phi=9, a/2=0{,}05)} = 1{,}833$$

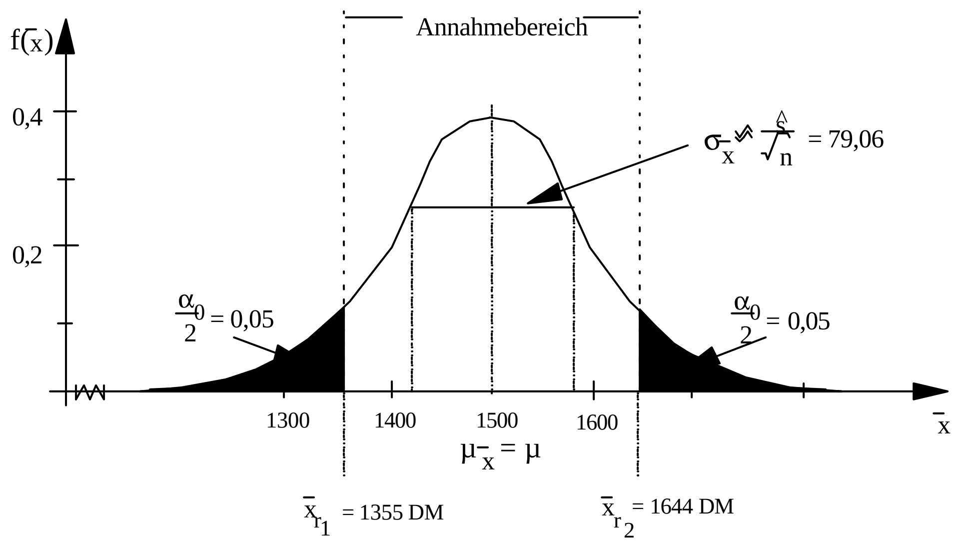
Einsetzen bei $a_0 = 0{,}1$:

$$\bar{x}_{r_1,r_2} = 1500 \pm 1{,}833 \cdot \frac{250}{\sqrt{10}} = 1500 \pm 1{,}833 \cdot 79{,}06$$

$$\bar{x}_{r_1} = 1355{,}09 DM$$

$$\bar{x}_{r_2} = 1644{,}91 DM$$

Graphische Darstellung des Annahmebereiches auf der Basis einer t-Verteilung bei a_0 = 0,10



Bei α_0 = 0,1 liegt der Annahmebereich der Verteilung zwischen 1355,09DM und 1644,91DM. Der Stichprobenwert von 1600DM führt also dazu, daß die Hypothese mit 95%iger Sicherheit angenommen werden kann, weil er im Annahmebereich liegt.

b) n = 550 $\bar{x} = 1550$DM $\hat{s}$ = 220DM

- σ ist unbekannt !
- $n \geq 31$

$$\frac{\bar{x} - m}{\frac{\hat{s}}{\sqrt{n}}}$$ ist normal-verteilt

- Endlichkeitsfaktor anwenden, weil $\frac{n}{N} = \frac{550}{10000} = 0,055 > 0,05$

$$\Rightarrow \quad \bar{x}_{r_1,r_2} = m_0 \pm Z_{\alpha_0/2} \cdot \frac{\hat{s}}{\sqrt{n}} \cdot \sqrt{\frac{N-n}{N-1}}$$

$\alpha_0 = 0{,}05 \Rightarrow Z_{\alpha_0/2} = 1{,}96$

Einsetzen bei $\alpha_0 = 0{,}05$:

$$\bar{x}_{r_1,r_2} = 1500 \pm 1{,}96 \cdot \frac{220}{\sqrt{550}} \cdot \sqrt{\frac{(10000-550)}{(10000-1)}} = 1500 \pm 1{,}96 \cdot 9{,}32 \cdot 0{,}945 = 1500 \pm 17{,}375 \Rightarrow$$

$\bar{x}_{r_1} = 1482{,}62 DM$

$\bar{x}_{r_2} = 1517{,}38 DM$

Bei α = 0,05 liegt die Stichprobe außerhalb des Annahmebereichs. Die Hypothese ist zurückzuweisen.

Aufgabe 33) „Hypothesentest für Anteilwerte“

Vorbemerkung
Die Grundgesamtheit wird in zwei Gruppen geteilt. A-Wähler und Nicht-A-Wähler. Es liegen also dichotome Merkmale vor; N = 50.000: die einzelnen Stichprobenelemente werden unabhängig voneinander gezogen; die Auswahlwahrscheinlichkeiten bleiben bei $\frac{n}{N} = \frac{15}{50000}$ annähernd konstant.
Es liegt also ein Bernoulli-Experiment vor $\Rightarrow$ Binomialverteilung.

Problem: Wird Partei A die 5%-Hürde $(p = 5\% = 0{,}05)$ schaffen?

An dieser Aufgabenstellung läßt sich die Problematik der Hypothesenformulierung darstellen.

Mögliche Hypothesen:
$H_0\!: p \geq p_0 = 0{,}05$
$H_0\!: p < p_0 = 0{,}05$
$H_0\!: \pi = \pi_0 = 0{,}05$

Mögliche Fehler:
α-Fehler: Ablehnung der richtigen Hypothese
β-Fehler: Annahme der falschen Hypothese.

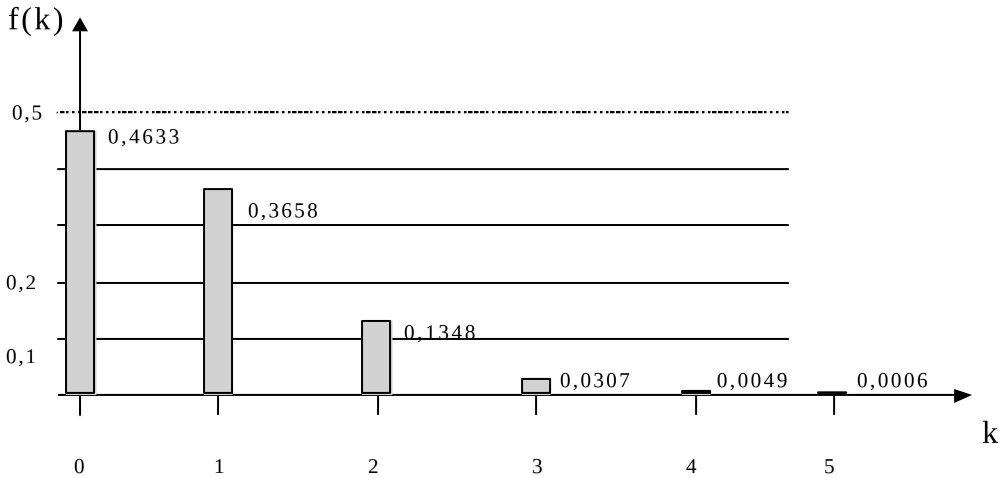
a) Hypothesentest auf der Basis von: n = 15 k = 0 a = 10% = 0,1

Prüfung der Approximationsmöglichkeit für die Normalverteilung $N_{(0,1)}$:
$VAR = n \cdot p \cdot (1-p) = 15 \cdot 0{,}05 \cdot 0{,}95 = 0{,}7125 < 9 \quad \Rightarrow$ keine Approximation

a1) In unserem Beispiel könnte das Interesse der Partei darin bestehen, in der Öffentlichkeit weiterhin die Hoffnung zu verbreiten, über die 5%-Hürde zu gelangen. Dies kann sie, wenn die Hypothese, $H_0\!: p \geq p_0 = 0{,}05$, nicht abgelehnt wird. Der α-Fehler bestünde dann darin, daß die Hypothese aufgrund einer StPr. mit einem Anteils-Wert deutlich unter 5% abzulehnen wäre, obwohl die Partei in der GG über die 5%-Hürde käme.

Tabellenwerte für k = 0 und folgende: vergleiche Graphik:

$B_{(n;k;p)} = B_{(15;0;0{,}05))} = 0{,}4633 > a \ (= 0{,}1)$



Aus der Graphik wird ersichtlich, daß ein linksseitiger Ablehnungsbereich bei gegbenem Signifikanzniveau nicht definiert werden kann.

⇒ 46% der Stichprobe aus der entnommenen GG mit $p \geq 0{,}05$ haben 0 Personen, die A wählen würden. Der Wert liegt oberhalb der Wahrscheinlichkeit des Ablehnungsbereiches, also kann die Hypothese weder abgelehnt noch angenommen werden.

a2) Es könnte aber auch das Interesse der Partei sein, ein möglichst ungeschminktes Bild der eigenen Chancen zu erhalten. Dazu wäre zu prüfen, ob die Hypothese H_0**:** $p \leq p_0 = 0{,}05$ abgelehnt werden kann. Der α-Fehler wäre dann die Ablehnung der Hypothese (die 5% werden nicht geschafft), obwohl diese für die GG zutrifft.

⇒ **Welche Hypothese zu wählen ist, hängt davon ab, wo die Folgewirkungen mit dem größere Risiko verbunden sind. Die Hypothese ist grundsätzlich so zu formulieren, daß mit dem** a **-Fehler das Risiko kontrolliert werden kann.**

Deshalb ist der Partei zumindest für die interne Information zu raten, die H_0**:** $p \leq p_0 = 0{,}05$ zu testen. Die Hypothese führt bei $P(k \geq 3) = 0{,}0362$ bzw. $P(k \geq 2) = 0{,}1710$ zu einem rechtsseitigen Ablehnungsbereich mit einem $a_0 = 0{,}0362$ bzw. $a_0 = 0{,}1710$. $a_0 = 0{,}1$ ist also direkt nicht realisierbar.

Da aber k = 0 ist, kann auch diese Hypothese nicht abgelehnt werden.

⇒ **Es ist unmöglich die beiden einander widersprechenden Hypothesen abzu-**

lehnen, da n = 15 zu klein ist.

Rat an die Partei: „Stichprobengröße erhöhen“

b) Erhöhung der Stichprobengröße

Vorbemerkung: Da die Bestimmung der Stichprobengröße unter Vorgabe des maximalen Stichprobenfehlers sachlich erst in Kap. 8.2.3 behandelt wird, soll hier nur geklärt werden wie groß die StPr. sein muß, damit der Anteilswert der StPr.normalverteilt ist. (Eine sich auf die Aufgabenformulierung im Buch beziehende Musterlösung wird unter Aufgabe 41 am ende des Textes behandelt.)

n soll so groß werden, daß eine Berechnung über die Normalverteilung erfolgen kann. Hierzu muß die Varianz ≥ 9 sein.

$$n \cdot p \cdot (1-p) \geq 9 \Rightarrow n \geq \frac{9}{(p \cdot (1-p))} \Rightarrow n \geq \frac{9}{0{,}05 \cdot 0{,}95} \Rightarrow n \geq 190$$

c) Hypothesentest auf der Basis von: n = 2000 k = 80 a = **0,01/0,05**

1. Es findet die Normalverteilung Anwendung, weil $n \cdot p \cdot (1-p) > 9$ (=2000) ist.
2. Der Endlichkeitsfaktor braucht nicht berücksichtigt zu werden, weil $\frac{n}{N} = \frac{2000}{50000} = 0{,}04 < 0{,}05$.

Die Frage lautet: „Überwindet Partei A die 5%-Hürde?“

c 1) Testen wir zuerst die Hypothese $H_0 : p \geq p_0$

Dieses ist die für die Partei in der Aussendarstellung bequemste Hypothese. Der linksseitige Test hat folgendes Ergebnis:
Die Formel für die Grenze des Annahmebereichs lautet:

$$p_r = p_0 - Z_{a_0} \cdot \sqrt{\frac{p_0 \cdot (1-p_0)}{n}}$$

Aus $Z_{(a=0{,}05)} = 1{,}65$ bzw. $Z_{(a=0{,}01)} = 2{,}33$ folgt:

$$a_0 = 0{,}05 \Rightarrow p_r = 0{,}05 - 1{,}65 \cdot \sqrt{\frac{0{,}05 \cdot 0{,}95}{2000}} = 0{,}042$$

$$a_0 = 0{,}01 \Rightarrow p_r = 0{,}05 - 2{,}33 \cdot \sqrt{\frac{0{,}05 \cdot 0{,}95}{2000}} = 0{,}0386$$

Vergleichen des empirischen Wertes mit dem Testmodell:

$$p = \frac{k}{n} = \frac{80}{2000} = 0{,}04.$$

Durch Erhöhung des Signifikanzniveaus von 5% auf 1% könnte hier also die Hypothese sogar noch gerettet werden, obwohl der beobachtete Stimmenanteil nur bei 4% liegt.. Ob dies die Partei beruhigt, ist eine andere Frage.

Das größte Risiko besteht allerdings für die Partei darin, daß sie die Hypothese aus c 1) annimmt, in der Wahl aber nicht über die 5%-Hürde gelangt. Diesem Risiko der Annahme der möglicherweise falschen Hypothese aus c 1) entspricht unserer Definition nach der b -Fehler. Da nur der a –Fehler von uns festlegbar ist, ist die Hypothese so umzuformulieren, dass, der a -Fehler dieses größte Risiko, fälschlicherweise von einem Stimmenanteil von über 5% auszugehen, beinhaltet. Dazu muß ein rechtsseitiger Test durchgeführt werden:

c 2) Test der Hypothese $H_0 : p \leq p_0$

$$p_r = p_0 + Z_{a_0} \cdot \sqrt{\frac{p_0 \cdot (1-p_0)}{n}}$$

Aus $Z_{(a=0{,}05)} = 1{,}65$ bzw. $Z_{(a=0{,}01)} = 2{,}33$ folgt:

$$a_0 = 0{,}05 \Rightarrow p_r = 0{,}05 + 1{,}65 \cdot \sqrt{\frac{0{,}05 \cdot 0{,}95}{2000}} = 0{,}058$$

$$a_0 = 0{,}01 \Rightarrow p_r = 0{,}05 + 2{,}33 \cdot \sqrt{\frac{0{,}05 \cdot 0{,}95}{2000}} = 0{,}0614$$

Da es sich um einen rechtsseitigen Test handelt, liegt das beobachtete p (=0,04) sowohl bei $\alpha = 0{,}01$ wie bei $\alpha = 0{,}05$ im Annahmebereich. Die Hypothese, die Partei schafft den Einzug ins Rathaus nicht, muß angenommen werden. Dies sollte für die Partei die Konsequenz haben, ihren Wahlkampf entsprechen zu verändern.

d) Wenn Partei C behaupten können möchte, daß A nicht in den Stadtrat einzieht, dann müßte die entsprechende Hypothese lauten:
$H_0 : p \leq p_0 = 0{,}05$, also: „A erreicht weniger als 5%“.

In diesem Fall handelt sich um den gleichen rechtsseitigen Test wie unter c 2).
Formel:

$$p_r = p_0 + Z_{a_0} \cdot \sqrt{\frac{p_0 \cdot (1-p_0)}{n}}$$

mit: $p_{(a=0{,}05)} = 0{,}0580$ und

$p_{(a=0,01)} = 0{,}0614$ (vgl. c 2))

Der sog. „Zurückweisungspunkt“ für die Hypothese ergibt sich jetzt aus:

$$p_r = \frac{k_r}{n} \Rightarrow k_r = p_r \cdot n$$

woraus der „Zurückweisungspunkt“ auf der Basis absoluter Anteile folgt:

$$\Rightarrow \begin{array}{l} k_{r_{(a=0,05)}} = 0{,}0580 \cdot 2000 = 116 \approx 116\,Personen \\ k_{r_{(a=0,01)}} = 0{,}0614 \cdot 2000 = 122{,}8 \approx 123\,Personen \end{array}$$

Sollten also weniger als 116 bzw. 123 Personen in einer Stichprobe mit n = 2000 sagen, sie würden A wählen wollen, so wäre die Hypothese „A verfehlt die 5%-Hürde“ anzunehmen.

e)

- Es ist bei einer seriösen Darstellung von Umfrageergebnisse also auf jeden Fall die Art der Hypothesenstellung zu hinterfragen.
- Ablehnung der Hypothese, die die negativsten Folgen nachsich ziehen würde, auf hohem Signifikanzniveau führt zu hoher Sicherheit.

Aufgabe 32 c)

Überprüfung der Standardabweichung der Stichprobenergebnisse aus Aufgabe 32 a und b.

c1) Zuerst die Konstellation aus Aufgabe 32 a)

n = 10 $\quad \overline{x} = 1600\,DM \quad \hat{s} = 250\,DM \quad a_0 = 0{,}1$

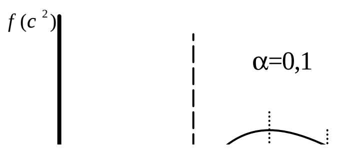
$n \leq 30 \quad s_0 = 200\,DM$

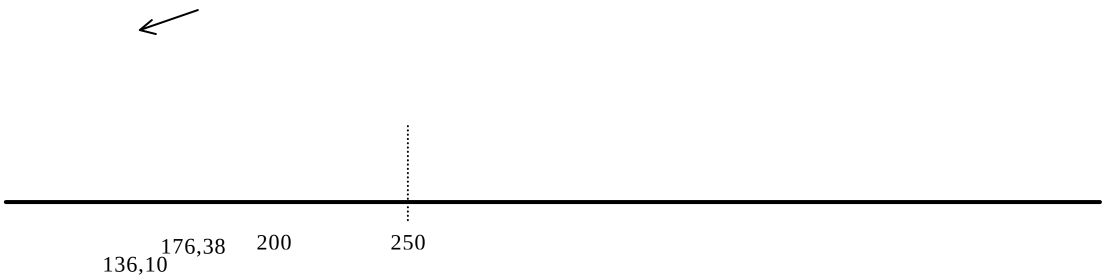
Wunsch: $s \leq 200$

Hypothese: $H_0 : s \geq s_0$ mit der Hoffnung, diese Hypothese ablehnen zu können.

In diesem linksseitigen Test läßt sich mittels des alpha-Fehlers kontrollieren, daß nicht vorschnell die Streuung des Einkommens als ausreichend gering eingeschätzt wird, da nur Stichproben zur Ablehnung der Hypothese führen, die eine sehr geringe Streuung des Einkommens aufweisen.

Graphische Darstellung:





Mit einem SN von 0,1 wird im linksseitiger Test, also mit: $c^{2}_{1-a_0} = 4{,}168$ die Nullhypothese angenommen mit:

$$P\left(c^{2}_{1-a_0} \le \frac{s_0^{2}}{n-1}\hat{s}^{2} \right) = 1 - a_0$$

$$P\left(\frac{s_0^{2}}{n-1} \cdot c^{2}_{1-a_0} \le \hat{s}^{2} \right) = 1 - a_0$$

$$P\left(\frac{200^{2}}{9} \cdot 4{,}168 \le \hat{s}^{2} \right) = 0{,}9$$

$$P\left(\hat{s}^{2}_{krit.} \le \hat{s}^{2} \right) = 0{,}9$$

$$\Rightarrow \hat{s}_{krit.} = 136{,}10\,DM \;\le 220\,DM$$

Nach dieser Rechnung liegt die Stichprobe des Jungakademikers im Annahmebereich. Die Hypothese, daß das Einkommen um **mehr** als 200 DM streut, kann also **nicht** abgelehnt werden. Würde man den a -Fehler, also das SN verkleinern, läge der für die Standardabweichung ermittelte Wert von 250 DM noch deutlicher im Annahmebereich.

c2) Für die Konstellation aus Aufgabe b) ergeben sich folgende Werte:

n = 550 $\quad \bar{x} = 1550$DM $\quad \hat{s}$ = 220DM

n > 30 $\quad a_0 = 0{,}05$

Hypothese (s.o.): $H_0 : s \ge s_0$

$Z_{a_0} = -1{,}65$

$$P\left(-Z_{a_0} \leq \sqrt{\frac{2f}{s_0^{\;2}}\hat{s}^2} - \sqrt{2f-1}\right) = 1 - a_0$$

$$P\left(\frac{\sqrt{2f-1} - Z_{a_0}}{\sqrt{2f}} \cdot s_0 \leq \hat{s}\right) = 1 - a_0$$

$$P\left(\frac{\sqrt{2 \cdot 549 - 1} - 1{,}65}{\sqrt{2 \cdot 549}} \cdot 200 \leq \hat{s}\right) = 0{,}95$$

$$P(\hat{s}_{krit.} \leq \hat{s}) = 0{,}95$$

$$\Rightarrow \hat{s}_{krit.} = 189{,}95 \leq 220\,DM$$

Auch diese Stichprobe liefert also einen Wert im Annahmebereich. Wieder kann die Hypothese, daß das Einkommen **mehr** als 200 DM streut **nicht** abgelehnt werden. Der Supermarkt sollte also nicht eröffnet werden.

Aufgabe 34:

a) gegeben: σ^2=0,01; ϕ=19
gesucht: P(0,005325< $\hat{s}^2$<0,01587)=?

$$c_u^2 = \frac{f}{s^2} \cdot \hat{s}_u^2 = \frac{19}{0{,}01} \cdot 0{,}005325 = 10{,}1175$$

$$c_o^2 = \frac{f}{s^2} \cdot \hat{s}_o^2 = \frac{19}{0{,}01} \cdot 0{,}01587 = 30{,}153$$

P(10,1175<χ^2<30,153)=0,9

Die Wahrscheinlichkeit, daß die Varianz des Durchmessers der Wellen im genannten Bereich liegen, beträgt 90%.

Aufgabe 35:

c^2-Anpassungstest

Die Nullhypothese zum Anpassungstest lautet immer:

H_0: „Die empirische Verteilung läßt sich durch die theoretische Verteilung hinreichend gut darstellen.“, d.h. $H_0: \; c_o^2 = 0$

Die empirischen Werte, also die erfragten Mietausgaben werden hierzu als beobachtete Werte f_b.betrachtet. Die errechneten Werte laut der Normalverteilung mit $m = 175DM$ und $s = 35DM$ sind die erwarteten Werte f_e.

Formel für c^2:

$$c^2 = \Sigma \frac{\left(f_b - f_e\right)^2}{f_e}$$

Die Rechnung wird mit absoluten Häufigkeiten und nicht mit Prozentwerten durchgeführt:

Tabelle: empirische und theoretisch erwartete Mietausgaben Oldenburger Studierender

Mietausgaben von... bis unter...	beobachtete Häufigkeiten f_b	erwartete Häufigkeiten f_e	$\left(f_b - f_e\right)$	$\frac{\left(f_b - f_e\right)^2}{f_e}$
0-130	17	19,71	-2,71	0,37
130-150	29	28,07	+0,93	0,03
150-170	40	41,10	-1,10	0,03
170-190	45	44,41	+0,59	0,01
190-210	29	34,99	-5,99	1,03
210-230	17	20,09	-3,09	0,48
230-∞	23	11,64	+11,36	11,09
Σ	200	200	-	$c^2 = 13{,}03$

Die Verteilung hat (k-1) Freiheitsgrade, also FG = 7-1 = 6
Wenn $c^2 = 0$ wäre, würden die empirischen Werte die gleichen sein, wie die theoretischen Werte, also eine optimale Abbildung der empirischen Verteilung wäre möglich. Deshalb ist der
c^2-Anpassungstest immer ein rechtsseitiger Test.
Die notwendige Voraussetzung, daß alle Klassen der Indifferenztabelle erwartete Häufigkeiten ≥ 5 aufweisen, ist erfüllt.

Bei a_0 = 0,05 und ϕ = 6 ergibt sich für c^2 ein Wert von 12,592 und bei
a_0 = 0,01 und ϕ = 6 ergibt sich für c^2 ein Wert von 16,812

Die Hypothese wird also bei einem Signifikanzniveau von 0,05 abgelehnt und bei einem Signifikanzniveau von 0,01 angenommen. Um sich nicht auf eine unsichere Hypothese zu stützen, empfiehlt es sich ein Signifikanzniveau von 1%.

Aufgabe 36-1(S. 330)**:**

c^2-Unabhängigkeitstest

Die Hypothese beim „Unabhängigkeitstest“ lautet immer:

H_0: „Die Ereignisse sind voneinander unabhängig“ (hier konkret: „Partei- und Ver
bandszugehörigkeit sind voneinander unabhängig“),
also wieder wie beim Anpassunmgstest:

$H_0: \quad c_o^2 = 0$

a) Berechnung von c^2:

Kontingenztabelle (f_b):

Partei	Arbeitnehmer-organisation	Industrie- und Arbeitgeber-organisation	Mittelstands-vereinigungen	Σ
CDU/CSU	19	15	44	**78**
SPD	43	0	1	**44**
Σ	**62**	**15**	**45**	**122**

Die Tabelle hat $(z-1)\cdot(s-1)=(3-1)\cdot(2-1)=2$ FG. Daraus folgt für die Indifferenztabelle, daß 2 Werte wie folgt berechnet werden müssen, der Rest ergibt sich von selbst:

$$f_{1,1}=\frac{78\cdot 62}{122}=39{,}64$$
$$f_{1,2}=\frac{78\cdot 15}{122}=9{,}59$$

Indifferenztabelle (f_e):

Partei	Arbeitnehmer-organisation	Industrie- und Arbeitgeber-organisation	Mittelstands-vereinigungen	Σ
CDU/CSU	**39,64**	**9,59**	28,77	**78**
SPD	22,36	5,41	16,23	**44**
Σ	**62**	**15**	**45**	**122**

Da alle Tabellenwerte der Indifferenztabelle $\geq$ 5 sind, können wir die Indifferenztabelle ohne weitere Zusammensetzung der Berechnung von c^2 zugrunde legen:

$$c^2 = \Sigma \frac{\left(f_b - f_e\right)^2}{f_e} = 10{,}75 + 3{,}05 + 8{,}06 + 19{,}05 + 15{,}41 + 14{,}29 = 60{,}61$$

b) Die Tabelle hat, wie erwähnt, 2 FG, also ergeben sich die c^2-Grenzwerte wie folgt:
bei α = 0,01 und 2 FG ist c^2 = 9,210
Der c^2-Wert von 60,61 liegt deutlich im Ablehnungsbereich. Die Hypothese wird deshalb abgelehnt. Partei- und Verbandszugehörigkeit sind voneinander abhängig.