

Dipl.-Kfm. Sascha Steinmann  
Universität Siegen  
Lehrstuhl für Marketing  
steinmann@marketing.uni-siegen.de

### Sommersemester 2010

### Marktforschung

#### Übungsaufgaben zur Varianzanalyse:

Die SAFTGENUSS GmbH plant die Einführung eines neuen Fruchtsaftes. Im Rahmen der Neuproduktentwicklung sollen zwei verschiedene Süßigkeitsgrade und drei verschiedene Farben getestet werden.

- a) Entwickeln und beschreiben Sie kurz ein empirisches Vorgehen, um den Einfluss von zwei Süßigkeitsgraden (niedrig vs. hoch) und drei Farben (Gelb-Orange vs. Orange vs. Blut-Orange) auf den Umsatz des Fruchtsaftes zu testen. Begründen Sie Ihren Vorschlag. Gehen Sie dabei auf Stärken und Schwächen des vorgeschlagenen Vorgehens ein.
- b) Eine Restaurantkette hat sich bereit erklärt, im Rahmen eines Experiments mit faktoriellem Design und Stichprobenumfang  $N = 18$  die entsprechenden Säfte über einen Zeitraum von 2 Monaten ihren Kunden anzubieten. Als Ergebnis wurden folgende Umsätze (in Euro) erzielt.

Farben (Index g)	Süßigkeitsgrad (Index h)		$\Sigma$
	Niedrig	Hoch	
Gelb-Orange	700	910	4920
	750	900	
	790	870	
Orange	850	940	5500
	880	970	
	930	930	
Blut-Orange	770	870	5030
	810	900	
	780	900	
$\Sigma$	7260	8190	15450

Diese Daten sollen per **mehrfaktorieller Varianzanalyse** (hier: zwei Faktoren) ausgewertet werden. Welche Voraussetzungen müssen für die Anwendung dieses Verfahrens erfüllt werden? Nennen Sie mindestens drei.

- c) Erklären Sie kurz das Grundprinzip der Varianzanalyse. Erklären Sie dazu vor allem, was das Prinzip der Streuungszerlegung ist, was Interaktionseffekte sind und aus welchen Größen sich der empirische F-Wert ergibt.

### **Ergänzung: Rechenaufgabe**

Führen Sie eine zweifaktorielle Varianzanalyse – basierend auf den Daten in Aufgabenteil b – mittels F-Test bei einem Signifikanzniveau von  $\alpha = 0,05$  durch!

### **Lösungsskizze:**

#### **a)**

- quasi-experimentelles Design: Auswahl von 6 möglichst identischen Supermärkten
- in jedem Supermarkt wird ein anderer Saft verkauft
- Erfassung des Umsatzes an vier aufeinander folgenden Wochen
- Stärke: z.B. Untersucht wird tatsächliches Kaufverhalten, Faktorabstufungen werden unabhängig voneinander realisiert, realistisches Kauf-Umfeld (kein Labor)
- Schwächen: z.B. keine zufällige Zuordnung der Probanden auf die Faktorabstufungen, nicht alle Umweltbedingungen können zwischen den Faktorstufen konstant gehalten werden, keine Information über den Grund des Kaufs

#### **b)**

- Formulierung einer Hypothese über den Wirkungszusammenhang der unabhängigen und der abhängigen Variablen.
- unabhängige Variablen auf nominalen, abhängige müssen auf metrischen Skalenniveau erhoben werden.
- die Faktoren müssen sich eindeutig unterscheiden (z.B. Verpackung und Markierung werden vom Käufer häufig gemeinsam wahrgenommen, somit wäre für diese Faktoren eine eindeutige Unterscheidung nicht gewährleistet).
- die Restgrößen wirken sich bis auf zufällige Schwankungen in allen Stichprobenzellen gleich aus (sog. Varianzhomogenität).
- die Werte in der Grundgesamtheit sind normalverteilt.
- die Additivität der Einflussgrößen, d.h. der Einfluss eines Faktors auf die Ergebnisvariable ist unabhängig vom Einfluss weiterer Faktoren oder auch Restgrößen.
- Strukturgleichheit, d.h. die in die Untersuchung gelangten Teilstichproben haben die gleiche Struktur der absatzbeeinflussenden Größen wie die Grundgesamtheit

#### **c)**

- Grundprinzip der Varianzzerlegung: Die gesamte Streuung/Varianz der Messwerte wird in zwei Varianzquellen zerlegt: Gesamtabweichung = durch die Faktorstufen erklärte Abweichung + durch die Interaktion der Faktoren erklärte Abweichung + nicht erklärte Abweichung
- Empirischer F-Wert: ergibt sich aus dem Verhältnis der erklärten Abweichung und der nicht erklärten Abweichung, ist umso größer, je mehr Varianz auf die Faktoren und deren Interaktion zurückzuführen ist

- Interaktionseffekte:

Die Wirkung einer UV auf die AV wird durch die Wirkung einer anderen UV beeinflusst

### Lösung: Ergänzung: Rechenaufgabe

Gegeben:  $G = 3$ ;  $H = 2$ ;  $K = 3$  (Anzahl an Beobachtungen pro Zelle);  $N = G * H * K = 18$

$$\begin{array}{ll}
 1. H_0^\alpha : \alpha_1 = \alpha_2 = \alpha_3 = 0 & H_1^\alpha : \text{min. ein } \alpha \neq 0 \\
 H_0^\beta : \beta_1 = \beta_2 = 0 & H_1^\beta : \text{min. ein } \beta \neq 0 \\
 H_0^{\alpha\beta} : \alpha\beta_{11} = \dots = \alpha\beta_{gh} = \dots = \alpha\beta_{32} = 0 & H_1^{\alpha\beta} : \text{min. ein } \alpha\beta \neq 0
 \end{array}$$

2. gegeben:  $\alpha = 0,05$

Entscheidungsregel: Lehne  $H_0$  ab, wenn gilt  $F_{\text{emp}} > F_{\text{tab}}$  (für alle  $H_0$ )

Tabellierte F-Werte:

$$\begin{array}{l}
 \text{Faktor 1 (Farbe): } F_{\text{tab}}^\alpha = F_{(1-\alpha, df_1, df_2)}^\alpha = F_{(0,95, 2, 12)}^\alpha = 3,89 \\
 \text{Faktor 2 (Süße): } F_{\text{tab}}^\beta = F_{(1-\alpha, df_1, df_2)}^\beta = F_{(0,95, 1, 12)}^\beta = 4,75 \\
 \text{Interaktion : } F_{\text{tab}}^{\alpha\beta} = F_{(1-\alpha, df_1, df_2)}^{\alpha\beta} = F_{(0,95, 2, 12)}^{\alpha\beta} = 3,89
 \end{array}$$

3. Berechnung der Mittelwerte:

$$\begin{array}{l}
 \bar{y}_{g=1} = 820 \quad \bar{y}_{g=2} = 916,66 \quad \bar{y}_{g=3} = 838,33 \\
 \bar{y}_{h=1} = 806,66 \quad \bar{y}_{h=2} = 910 \\
 \bar{y} = 858,33
 \end{array}$$

$$\begin{array}{l}
 \bar{y}_{11} = 746,66 \quad \bar{y}_{12} = 893,33 \\
 \bar{y}_{21} = 886,66 \quad \bar{y}_{22} = 946,66 \\
 \bar{y}_{31} = 786,66 \quad \bar{y}_{32} = 890
 \end{array}$$

Berechnung der Schätzwerte für die einzelnen Zellen:

$$\begin{array}{l}
 \hat{y}_{gh} = \bar{y}_g + \bar{y}_h - \bar{y} \\
 \hat{y}_{11} = 768,34 \quad \hat{y}_{12} = 871,67 \\
 \hat{y}_{21} = 865,01 \quad \hat{y}_{22} = 968,34 \\
 \hat{y}_{31} = 786,67 \quad \hat{y}_{32} = 890
 \end{array}$$

4. Berechnung der quadrierten Abweichungen:

$$SS_a = H \cdot K \sum_{g=1}^G (\bar{y}_g - \bar{y})^2$$

$$= 2 \cdot 3 \cdot 5272,4 = 31636,467$$

$$SS_b = G \cdot K \sum_{h=1}^H (\bar{y}_h - \bar{y})^2$$

$$= 48046,9$$

$$SS_{a \times b} = K \sum_{g=1}^G \sum_{h=1}^H (\bar{y}_{gh} - \hat{y}_{gh})^2$$

$$= 5632,467$$

$$SS_t = \sum_{g=1}^G \sum_{h=1}^H \sum_{k=1}^K (\bar{y}_{ghk} - \bar{y})^2$$

$$= 95850$$

$$SS_w = SS_t - SS_a - SS_b - SS_{a \times b} = 10534,166$$

5. Berechnung der mittleren quadrierten Abweichungen:

$$MS_a = \frac{SS_a}{G-1} = 15818,23$$

$$MS_b = \frac{SS_b}{H-1} = 48046,9$$

$$MS_{a \times b} = \frac{SS_{a \times b}}{(G-1)(H-1)} = 2816,23$$

$$MS_w = \frac{SS_w}{G \cdot H \cdot (K-1)} = 877,85$$

6. Berechnung der empirischen Prüfgröße:

$$F_{emp}^i = \frac{MS_i}{MS_w}$$

$$F_{emp}^\alpha = 18,02$$

$$F_{emp}^\beta = 54,325$$

$$F_{emp}^{\alpha\beta} = 3,208$$

7. Vergleich zwischen empirischen (berechneten) und tabellierten (theoretischen) F-Wert anhand der relevanten Entscheidungsregel:

$$F_{emp}^{\alpha} = 18,02 > 3,89 = F_{tab}^{\alpha}$$

$$F_{emp}^{\beta} = 54,325 > 4,75 = F_{tab}^{\beta}$$

$$F_{emp}^{\alpha\beta} = 3,208 < 3,89 = F_{tab}^{\alpha\beta}$$

Für die Faktoren „Farbe“ und „Süßigkeitsgrad“ wird  $H_0$  abgelehnt, d.h. sowohl die Farbe als auch der Süßigkeitsgrad haben einen Einfluss auf den Absatz. Die Interaktion zwischen beiden Faktoren ist hingegen nicht signifikant, d.h. die Faktoren hängen nicht voneinander ab.