Fachbereich 5
Lehrstuhl für Marketing
Univ.-Professor Dr. Hanna Schramm-Klein

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

Sommersemester 2010

Marktforschung

Übungsaufgaben zu den Themen 3-6 mit Lösungsskizzen

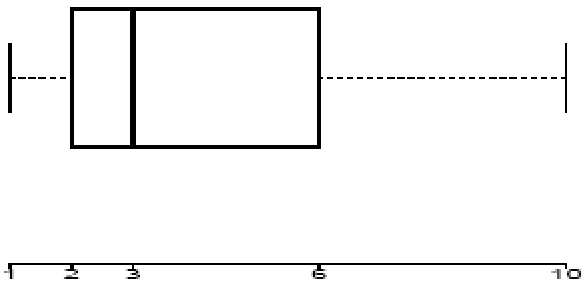
Aufgabe 1:

Geben Sie an, ob die folgenden Aussagen **richtig** oder **falsch** sind.

	Richtig	Falsch
1) Sind zwei metrisch skalierte Variablen unkorreliert (r=0), folgt daraus, dass sie unabhängig von einander sind, d.h. dass kein Zusammenhang zwischen ihnen besteht.	O	X
2) Extreme Beobachtungswerte haben einen großen Einfluss auf den Mittelwert einer Verteilung.	X	O
3) Ein Fehler erster Art wird begangen, wenn die Nullhypothese verworfen wird, obwohl sie richtig ist.	X	O
4) Eine Hypothese werde auf dem Signifikanzniveau 0.05 verworfen. Dann kann sie immer auch auf dem Signifikanzniveau 0.025 verworfen werden, wenn der kritische Bereich auf der richtigen Seite des Wertebereichs der Prüfgröße gewählt wurde.	O	X
5) Die wichtigsten Lageparameter zur Beschreibung einer Stichprobe sind Mittelwert, Median und Varianz.	O	X
6) Nichtparametrische Tests basieren auf Rangwerten, die den Originalmesswerten zugeordnet werden.	X	O
7) Die Standardabweichung ist ein geeignetes Streuungsmaß für metrisch skalierte Merkmale.	X	O
8) Assoziationsanalysen untersuchen gerichtete Abhängigkeiten.	O	X
9) Der Kontingenzkoeffizient ist eine geeignete Maßzahl, um Stärke und Richtung eines Zusammenhangs zu beschreiben.	O	X
10) Rangkorrelationskoeffizienten können Beziehungen zwischen zwei mindestens ordinal skalierten Merkmalen aufdecken.	X	O

Aufgabe 2:

Bei einer Untersuchung wurde der tägliche Fernsehkonsum in Stunden von 200 Probanden erhoben. Die Verteilung lässt sich mit Hilfe eines Boxplots wie folgt darstellen:



a) **Nennen** und **erläutern** Sie die Werte, die für die Erstellung eines Boxplots notwendig sind.
b) **Interpretieren** Sie diesen darstellten Boxplot und **zeichnen** Sie zusätzlich die **Spannweite** und den **Interquartilsabstand** ein. **Beurteilen** Sie visuell die **Schiefe** der Verteilung.

a)

Boxplot als Fünf-Punkte-Zusammenfassung:

- Minimum: kleinster Wert einer Verteilung
- 1. Quartil: Median der unteren Hälfte des geordneten Datensatzes
- Median: mittlere Wert in einer Reihe der nach der Größe geordneten Werte
- 3. Quartil: Median der oberen Hälfte des geordneten Datensatzes
- Maximum: größter Wert einer Verteilung

b)

- Die Quartile, das Minimum, Maximum sowie der Median teilen den Datensatz in vier gleiche Teile, wobei jeder dieser Teile etwa ein Viertel der Beobachtungswerte (also 50) enthält.
- keine Ausreißer im Datensatz vorhanden!
- man kann nicht erkennen, ob mehrere Gipfel in dem Datensatz vorliegen!
- Spannweite/Quartilsabstand als Indikator für die Streuung
- Die Lage der Medianlinie weist Asymmetrie auf. Da die Medianlinie in Richtung des 1. Quartils weist, liegt hier eine rechtsschiefe bzw. linkssteile Verteilung vor.

Aufgabe 3

Zur Beschreibung des Zusammenhangs zwischen metrisch skalierten und normalverteilten Variablen wird häufig der **Korrelationskoeffizient nach Bravais-Pearson** berechnet.

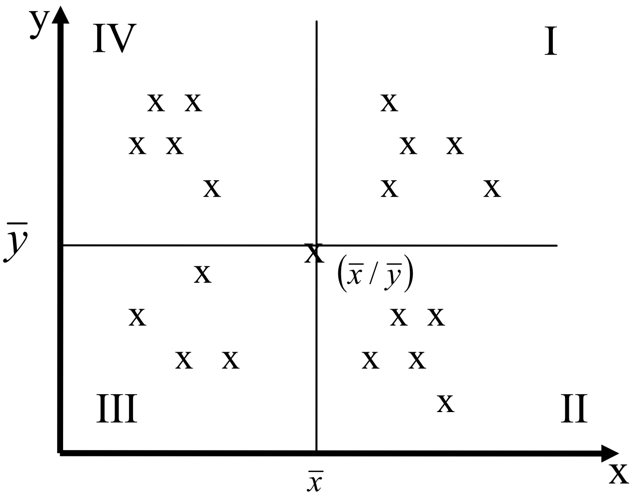
a) **Nennen** und **erläutern** Sie die Maße, die in die Berechnung des Korrelationskoeffizienten nach Bravais-Pearson eingehen.

b) **Erläutern** Sie den Korrelationskoeffizienten nach Bravais-Pearson. Berücksichtigen Sie hierbei v.a. den möglichen **Wertebereich** sowie unterschiedliche **Arten von Zusammenhängen**. Unterstützen Sie Ihre Ausführungen mit Hilfe **grafischer Darstellungen**.

a)

1. empirische Kovarianz

- durchschnittliche Summe von Abweichungsprodukten
- jeder Summand hat die Struktur $(x_i - \bar{x})(y_i - \bar{y})$
- $\bar{x}$ und $\bar{y}$ stellen den Schwerpunkt der Punktwolke dar
- aus der Darstellung ergeben sich unmittelbar die Vorzeichen der einzelnen Komponenten und des Produkts (in I und III = positiv, in II und IV = negativ)



- die Kovarianz gibt die Tendenz an, in welche Richtung die Merkmale variieren

  $s_{xy} = 1/n \sum_{i=1}^{n} (x_i - \bar{x})(y_i - \bar{y})$

  $s_{xy} > 0$ mit x steigt (tendenziell) auch y (und umgekehrt)

  $s_{xy} < 0$ hohe Werte der einen Zufallsvariablen gehen mit niedrigen Werten der anderen Zufallsvariablen einher

  $s_{xy} = 0$ x und y sind unabhängig
- Kovarianzen deuten (ggf.) auf lineare Abhängigkeiten hin. Sie sind von den Maßeinheiten der Merkmale abhängig!
- Wertebereich : $-\infty$ bis $+\infty$

2. Standardabweichung

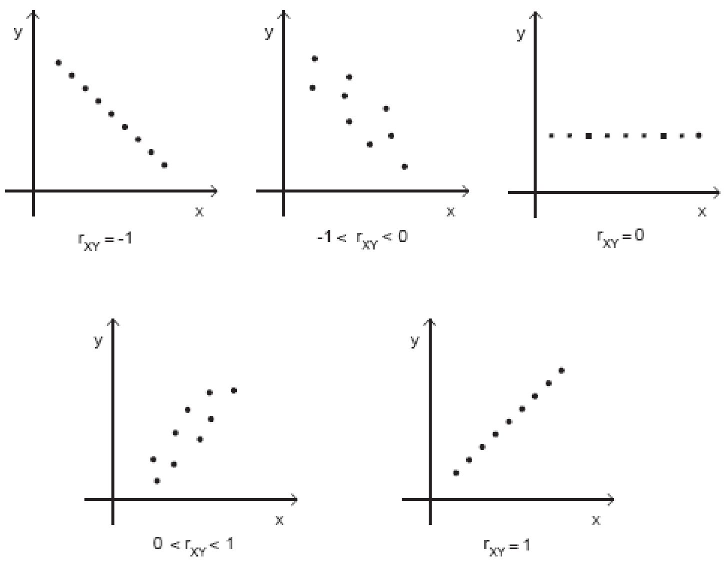
- Quadratwurzel aus der Varianz eines Datenbündels (Varianz = Summe der quadrierten Abweichungen der einzelnen Werte x_i eines Datenbündels vom Mittelwert, dividiert durch die Anzahl der Beobachtungen *n)*
- nur für metrische Daten anwendbar
- eignet sich zur Kennzeichnung von Fehlerintervallen um das arithmetische Mittel
- durch die Wurzelberechnung wird die Quadrierung der Abweichungen "rückgängig gemacht", so dass *s* die gleiche Maßeinheit hat wie die Datenwerte selbst

b)

- **Korrelationskoeffizienten nach Bravais-Pearson (Produkt-Moment-Korrelation) r_{xy}:** Division der Kovarianz durch die Standardabweichungen beider Merkmale (=Eliminierung der Streuung der einzelnen Verteilungen)

$$r_{xy} = \frac{\sum_{i=1}^{n}(x - \bar{x})(y - \bar{y})}{\sqrt{\sum_{i=1}^{n}(x_i - \bar{x})^2 \sum_{i=1}^{n}(y_i - \bar{y})^2}} = \frac{s_{xy}}{s_x s_y}$$

- Wertebereich von r_{xy} : -1 bis +1
- positive r_{xy} → die Merkmale variieren tendenziell in der gleichen Richtung
- negative r_{xy} → die Merkmale variieren tendenziell in entgegengesetzter Richtung
- $r_{xy} = 0$ → kein (linearer) Zusammenhang!
- z.B.



Aufgabe 4

Anlässlich der geplanten Einführung eines neuen Produktes wurde eine Befragung durchgeführt. Die Probanden wurden hierbei u.a. nach ihrer Schulbildung (1= Hauptschule, 2= Mittlere Reife, 3= Abitur) und der Einschätzung der Wichtigkeit, die das neue Produkt für sie hat, befragt. Die Einschätzung der Wichtigkeit erfolgte auf einer zehn Punkte umfassenden Skala (von 1= völlig unwichtig bis 10=sehr wichtig).

Die nachstehende Tabelle zeigt die Ergebnisse der Befragung. Hierbei wurde bereits eine Gruppenbildung hinsichtlich der Schulbildung vorgenommen.

Hauptschule	Mittlere Reife	Abitur
5	5	3
5	3	5
7	3	4
9	8	5
5	6	3
6	7	4
8	5	6
6	3	2
10	4	
	7	
	7	

a) Welchen **Test** würden Sie anwenden, um zu prüfen, ob der Unterschied zwischen den Werten signifikant ist (p<0,05)? **Begründen** Sie Ihre Antwort.

b) Formulieren Sie die dem Test zugrunde liegenden **ungerichteten Hypothesen** (Alternativ- und Nullhypothese).

c) **Überprüfen** Sie auf einem Signifikanzniveau von 0.05, ob der Unterschied zwischen den Werten signifikant ist. **Erläutern** Sie Ihr Vorgehen.

Hinweise:

$$U = R - \frac{n(n+1)}{2}$$

$$H = \frac{12}{n(n+1)} \sum_{i=1}^{k} \frac{T_i^2}{n_i} - 3*(n+1)$$ mit $df=k$-1 Freiheitsgraden

$$t = \frac{|\bar{d}|\sqrt{n}}{s}$$ mit $df=n$-1 Freiheitsgraden

(a)

- Vergleich von mehr als zwei unabhängigen Stichproben
- Werte weisen Ordinalskalenniveau auf
  → H-Test nach Kruskal und Wallis

(b)

- Nullhypothese: Die drei Kategorien der Schulbildung unterscheiden sich nicht bezüglich der Einschätzung der Wichtigkeit, die das neue Produkt für sie hat.
- Alternativhypothese: Die drei Kategorien der Schulbildung unterscheiden sich bezüglich der Einschätzung der Wichtigkeit, die das neue Produkt für sie hat.

(c)

- Bestimmung der gemeinsamen Rangreihe und Ermittlung der Rangsummen

Hauptschule		Mittlere Reife		Abitur	
Wert	**Rang**	Wert	**Rang**	Wert	**Rang**
5	13	5	13	3	4
5	13	3	4	5	13
7	22,5	3	4	4	8
9	27	8	25,5	5	13
5	13	6	18,5	3	4
6	18,5	7	22,5	4	8
8	25,5	5	13	6	18,5
6	18,5	3	4	2	1
10	28	4	8		
		7	22,5		
		7	22,5		
Summe	179	**Summe**	157,5	**Summe**	69,5

- Berechnung der Prüfgröße H des Kruskal-Wallis-Tests und der Freiheitsgrade mit Hilfe der folgenden Werte:

k=3
n_1=9
n_2=11
n_3=8
$n=28$
T_1=179
T_2=157,5
T_3=69,5

$$H = \frac{12}{28*29} * \left(\frac{179^2}{9} + \frac{157{,}5^2}{11} + \frac{69{,}5^2}{8} \right) - 3*29 = 7{,}862$$

mit df = 3-1= 2 Freiheitsgrade

- Vergleich des Prüfgrößenwertes mit dem kritischen Tabellenwert der Chiquadrat-Tabelle.

  Der kritische Tabellenwert lautet: 7,38. Da der berechnete Wert den tabellierten Wert übersteigt, liegt Signifikanz auf der betreffenden Stufe vor.
  Ergebnis: Die Nullhypothese kann damit verworfen werden.

t-Tabelle:

	Sicherheit in %;		
f	95%	99%	99, 9%
1	12, 71	63, 66	636, 62
2	4, 30	9, 92	31, 60
3	3, 18	5, 84	12, 92
4	2, 78	4, 60	8, 61
5	2, 57	4, 03	6, 86
6	2, 45	3, 71	5, 96
7	2, 37	3, 50	5, 41
8	2, 31	3, 36	5, 04
9	2, 26	3, 25	4, 78
10	2, 23	3, 17	4, 59
11	2, 20	3, 11	4, 44
12	2, 18	3, 06	4, 32
13	2, 16	3, 01	4, 22
14	2, 15	2, 98	4, 14
15	2, 13	2, 95	4, 07
16	2, 12	2, 92	4, 02
17	2, 11	2, 90	3, 96
18	2, 10	2, 88	3, 92
19	2, 09	2, 86	3, 88
20	2, 08	2, 85	3, 85
25	2, 060	2, 787	3, 725
30	2, 042	2, 750	3, 646
35	2, 030	2, 724	3, 592
40	2, 021	2, 704	3, 551
45	2, 014	2, 689	3, 521
50	2, 009	2, 678	3, 496
100	1, 984	2, 626	3, 390
200	1, 972	2, 601	3, 340
300	1, 969	2, 595	3, 328
400	1, 967	2, 590	3, 318
500	1, 965	2, 586	3, 310
600	1, 964	2, 585	3, 307
700	1, 963	2, 584	3, 304
800	1, 963	2, 583	3, 302
∞	1, 960	3, 576	3, 291

χ^2-Tabelle

Für 90% bis 99,9% Sicherheit:

f	90%	95%	97,5%	99%	99,5%	99,9%
1	2,71	3,84	5,02	6,63	7,88	10,83
2	4,61	5,99	7,38	9,21	10,60	13,82
3	6,25	7,81	9,35	11,35	12,64	16,27
4	7,78	9,49	11,14	13,28	14,86	18,47
5	9,24	11,07	12,83	15,09	16,75	20,52
6	10,64	12,59	14,45	16,81	18,55	22,46
7	12,02	14,07	16,01	18,48	20,78	24,32
8	13,36	15,51	17,53	20,09	21,96	26,13
9	14,68	16,92	19,02	21,67	23,59	27,88
10	15,99	18,31	20,48	23,21	25,19	29,59
11	17,28	19,68	21,92	24,72	26,76	31,26
12	18,55	21,03	23,34	26,22	28,30	32,91
13	19,81	22,36	24,74	27,69	29,82	34,53
14	21,06	23,68	26,12	29,14	31,32	36,12
15	22,31	25,00	27,49	30,58	32,80	37,70
16	23,54	26,30	28,85	32,00	34,27	39,25
17	24,77	27,59	30,19	33,41	35,72	40,79
19	25,99	28,87	31,53	34,81	37,16	42,31
19	27,20	30,14	32,85	36,19	38,58	43,82
20	28,41	31,41	34,17	37,57	40,00	45,31
21	29,62	32,67	35,48	38,93	41,40	46,80
22	30,81	33,92	36,78	40,29	42,80	48,27
23	32,01	35,17	38,08	41,64	44,18	49,73
24	33,20	36,42	39,36	42,98	45,56	51,18
25	34,38	37,65	40,65	44,31	46,93	52,62
26	35,56	38,89	41,92	45,64	48,29	54,05
27	36,74	40,11	43,19	46,96	49,64	55,48
28	37,92	41,34	44,46	48,28	50,99	56,89
29	39,09	42,56	45,72	49,59	52,34	58,30
30	40,26	43,77	46,98	50,89	53,67	59,70
31	41,42	44,99	48,23	52,19	55,00	61,10
32	42,59	46,19	49,49	53,49	56,33	62,49
33	43,75	47,40	50,73	54,78	57,65	63,87
34	44,90	48,60	51,97	56,06	58,96	65,25
35	46,06	49,80	53,20	57,34	60,27	66,62
36	47,21	51,00	54,44	58,62	61,58	67,98
37	48,36	52,19	55,67	59,89	62,88	69,34
38	49,51	53,38	56,90	61,16	64,18	70,70
39	50,66	54,57	58,12	62,43	65,48	72,05
40	51,81	55,76	59,34	63,69	66,77	73,40