

Prüfung zur Datenermittlung SS 2020*

Prüfer: Prof. Dr. J. Dovern

Arbeitszeit: 60 Minuten

Wahl der Aufgaben: Es sind alle Aufgaben zu bearbeiten. Es können insgesamt maximal 60 Punkte erreicht werden.

Hilfsmittel: Zugelassen ist ein Taschenrechner und eine selbstgeschriebene Formelsammlung im Umfang von 2 DIN A4-Seiten. Beim Taschenrechner dürfen jedoch keine Programme oder Programmteile verwendet werden, die nicht fest in den Taschenrechner eingebaut sind. Die Hilfsmittel sind mitzubringen.

Bewertung: Bewertet werden grundsätzlich nur Lösungen, die im Lösungsteil stehen und für die folgendes beachtet wird:

- Antworten sind immer zu begründen, es sei denn, es wird ausdrücklich keine Begründung gefordert.
- Der Lösungsweg muss aus einer Darstellung der einzelnen Rechenschritte ersichtlich sein. Geben Sie Lösungen grundsätzlich mit drei Nachkommastellen an, wenn nicht ausdrücklich anders gefordert.
- Wird die Bearbeitung einer Aufgabe unterbrochen und an anderer Stelle wieder aufgenommen, so ist an der Stelle, an der sie unterbrochen wurde, ein entsprechender Vermerk anzubringen.
- Unleserliche Aufgabenteile werden mit Null Punkten bewertet.

Viel Erfolg bei der Bearbeitung der Aufgaben!

*korrigierte Version

Aufgabe 1

Sie sind Marktforscherin bzw. Marktforscher bei einem Baumarktunternehmen. Das Unternehmen überlegt, in der Stadt ABC einen zusätzlichen Baumarkt zu eröffnen.

a) Sie werden gefragt, ob eine vorherige Marktforschung sinnvoll ist, wenn ein Zeithorizont von 10 Jahren zugrunde gelegt wird. Folgende Daten sind bekannt:

- Wird der zusätzliche Baumarkt erfolgreich, dann werden in den nächsten 10 Jahren zusätzliche Gewinne von 10 Mio. Euro erzielt. Ist er ein Flop, so entstehen im gleichen Zeitraum zusätzliche Verluste von 8 Mio. Euro.
- Eine Marktforschung kann den Erfolg nur mit einer gewissen Wahrscheinlichkeit vorhersagen. Wird ein Erfolg vorhergesagt, dann ist dennoch mit einer Wahrscheinlichkeit von 20% ein Flop zu erwarten. Wird dagegen ein Flop vorhergesagt, ist dennoch mit einer Wahrscheinlichkeit von 20% ein Erfolg zu erwarten.
- Ohne vorherige Marktforschung sind Erfolg und Flop mit je 50% gleich wahrscheinlich.
- Die Marktforschung kostet 100.000 Euro. Die Wahrscheinlichkeit, dass die Marktforschung Flop bzw. Erfolg prognostiziert wird mit je 50% beziffert.

Empfehlen Sie die Durchführung der Marktforschung? Rechnerische Begründung!
(6 Punkte)

Lösung:

Ohne Mafo: Erwartungswert Gewinn: $0,5 \cdot 10 \text{ Mio.} + 0,5 \cdot (-8 \text{ Mio.}) = 1 \text{ Mio.}$

Mit Mafo: Erwartungswert Gewinn:

$0,5 \cdot (0,8 \cdot 10 \text{ Mio.} + 0,2 \cdot (-8 \text{ Mio.})) + 0 \text{ Euro} - 0,1 \text{ Mio.} = 3,1 \text{ Mio.}$

Erwartungswert mit Mafo höher → durchführen.

b) Die Marktforschung soll durchgeführt werden. Dabei werden zwei Teilzielgruppen definiert:

1. Alle Haushalte, die in einem Umkreis von 20 Minuten Fahrzeit leben als Kernzielgruppe. Es gibt 40.000 Haushalte in der Kernzielgruppe.

2. Alle Haushalte, die in einem Umkreis ab 20 Minuten bis 45 Minuten Fahrzeit leben als erweiterte Zielgruppe. Es gibt 100.000 Haushalte der erweiterten Zielgruppe.

Es sollen nur Personen befragt werden, die in letzter Zeit Heimwerkerarbeiten durchgeführt haben.

b1) Warum ist die Definition „Personen, die in letzter Zeit Heimwerkerarbeiten durchgeführt haben“ nicht geeignet? Formulieren Sie eine mögliche geeignete Definition! (4 Punkte)

Lösung:

Keine klare Definition. Was heißt „in letzter Zeit“? Was sind „Heimwerkertätigkeiten“?

Mögliche Definition: Personen, die in ihrem Haushalt in den letzten 3 Monaten insgesamt mindestens 2 Stunden lang Heimwerkerarbeiten durchgeführt haben.

b2) In der Umfrage sollen u.a. folgende Themen abgefragt werden:

1. Image des eigenen Baumarktunternehmens und des im Ort bereits ansässigen Konkurrenzunternehmens.
2. Absicht, einen neuen Baumarkt zu besuchen, wenn ein solcher eröffnet wird.
3. Absicht, innerhalb des nächsten halben Jahres Renovierungen im oder am eigenen Einfamilienhaus durchzuführen.
4. Freude an der Durchführung verschiedener gärtnerischer Tätigkeiten.

b2-1) Bei welchen dieser Themen schlagen Sie vor, mit einer Multi-Item-Skala zu arbeiten, bei welchen mit einem Einzelitem? (4 Punkte)

Lösung:

- 1. und 4. sind komplexe Sachverhalte mit mehreren Aspekten → Multi-Item-Skala*
- 2. und 3. können mit einer Frage beantwortet werden → Einzelitem*

b2-2) Bei welchem Thema bzw. bei welchen Themen ist eine vorherige Filterfrage erforderlich? (2 Punkte)

Lösung:

- 1., da Baumarktunternehmen bekannt sein muss.*
- 3., da nicht jeder im eigenen Einfamilienhaus wohnt.*
- 4., da das nur beantworten kann, wer das schon getan hat, als i.d.R. einen Garten hat.*

c) Zum Themenkomplex „Freude an der Durchführung handwerklicher Tätigkeiten werden u.a. folgende Statements vorgeschlagen:

1. Heimwerken macht mir Freude
2. Ich freue mich gar nicht, wenn es im Haushalt nichts zum Heimwerken gibt
3. Handwerkliche Tätigkeiten im Haushalt gehen einfach und schnell
4. Handwerkliche Tätigkeiten im Haushalt machen mir schon ein bisschen Spaß

Welche dieser Statements sind ungeeignet und warum? (4 Punkte)

Lösung:

- 1. OK*
- 2. Schlecht, da doppelte Verneinung*
- 3. Schlecht da zwei Sachverhalte in einem Statement*
- 3. Schlecht da kein monotoner Zusammenhang mit Sachverhalt*

d) Sie haben 5 Statements für die Skala „Freude an der Durchführung handwerklicher Tätigkeiten“ entwickelt und in einem Pretest mit einer Likert-Skala von 1-5 abgefragt. Dabei ergab sich folgendes Ergebnis:

	Statements					
	1	2	3	4	5	Summe
Person 1	1	2	1	4	2	10
Person 2	3	3	2	5	4	17
Person 3	2	5	3	4	3	17
Person 4	4	5	4	5	5	23
Person 5	3	2	5	5	3	18
Person 6	3	2	4	4	5	18
Person 7	2	3	1	5	3	14
Summe	18	22	20	32	25	117
Varianz	0,816	1,551	2,122	0,245	1,102	13,633
Itemschwierigkeit		3,143	2,857	4,571	3,571	
Itemtrennschärfe		0,230	0,494	0,281	0,677	

d1) Berechnen Sie ein Maß für die Validität der Skala und beurteilen Sie das Ergebnis! (4 Punkte)

Lösung:

$$\text{Cronbach's Alpha} = 5/4 * (1 - (0,816 + 1,551 + 2,122 + 0,245 + 1,102) / 13,633) = 0,715$$

Das ist größer 0,7 und damit OK.

d2) Bestimmen Sie die Itemschwierigkeit und die Itemtrennschärfe für Statement 1! (4 Punkte)

Lösung:

$$\text{IS: } 18/7 = 2,57$$

IT: Korrelation des Statements mit Summe minus Statement:

1 9
3 14
2 15
4 19
3 15
3 15
2 12

Korrelation ergibt 0,911

d3) Welche der Statements 2 bis 5 würden Sie nicht für den endgültigen Fragebogen verwenden und warum? (4 Punkte)

Lösung:

*Itemschwierigkeit sollte zwischen 30 und 70% der Spannweite sein, also 2,2 bis 3,8
Wird verletzt von Item 4*

Itemtrennschärfe sollte sein >0,3. Wird verletzt von Item 2 und Item 4 → Items 2 und 4 sind nicht geeignet.

e) Die Befragung einer geschichtete Zufallsstichprobe mit je 500 Personen der Schicht „Kernzielgruppe“ und 500 Personen der Schicht „erweiterter Zielgruppe“ ergab u.a. folgendes Ergebnis: Anteil der Personen mit der Absicht, einen neuen

Baumarkt zu besuchen:
Kernzielgruppe: 60%
Erweiterte Zielgruppe: 42%

e1) Welche Schätzung ergibt sich für den Anteil in der Grundgesamtheit? (2 Punkte)

Lösung:

$$(0,6 \cdot 40 + 0,42 \cdot 100) / 140 = 0,471$$

e2) Der Marketingleiter behauptet, dass 50% aller Personen der Zielgruppe einen neuen Baumarkt besuchen wollen. Können Sie dies mit einer Sicherheitswahrscheinlichkeit von 95% ablehnen? (6 Punkte)

Hinweis: $z(92,5\%)=1,44$, $z(95\%)=1,64$; $z(97,5\%)=1,96$. Haben Sie keine Lösung bei Aufgabe e1, so rechnen Sie mit 47,3% (dies ist nicht die Lösung von e1).

Lösung:

$$\sigma^2(\bar{x}_{st}) = \sum_{h=1}^L W_h^2 \cdot \frac{\sigma_h^2}{n_h}$$

$= 0,286^2 \cdot 0,6 \cdot 0,4 / 500 + 0,714^2 \cdot 0,42 \cdot 0,58 / 500 = 0,000288 \rightarrow s(x) = 0,017 \rightarrow$
Vertrauensintervall: $[0,471 - 1,96 \cdot 0,017; 0,471 + 1,96 \cdot 0,017] = [0,438; 0,504] \rightarrow$
50% im Vertrauensintervall \rightarrow Hypothese kann nicht abgelehnt werden.

Aufgabe 2

Sie sind Marktforscher bzw. Marktforscherin bei einer bundesweit tätigen Bank.

a) Die Bank möchte das Anlageverhalten der privaten Haushalte in Deutschland kontinuierlich verfolgen. Sie bekommen dazu zwei Angebote.

Angebot 1:

Panel von 10.000 Haushalten repräsentativ. Das Panel hat eine Panelsterblichkeit von 24% pro Jahr.

Angebot 2:

Wellenbefragung von 18.000 Haushalten repräsentativ.

Die Bank weiß, dass das Anlageverhalten zwischen zwei Jahren mit $r = 0,68$ korreliert. Die Streuung des Anlageverhaltens bleibt weitgehend konstant.

a) Was haben eine Wellenbefragung und ein Panel gemeinsam und worin unterscheiden sie sich? (4 Punkte)

Lösung:

Gemeinsamkeit: Gleicher Sachverhalt wird zu regelmäßig wiederkehrenden Zeitpunkten in der gleichen Methode abgefragt. Unterschied: Panel hat gleiche, Wellenbefragung unterschiedliche Stichprobe.

b) Geben Sie je ein Beispiel für einen Sachverhalt, der nicht in einem Panel abgefragt werden sollte und einen, der für ein Panel geeignet ist und nehmen Sie Stellung dazu, ob das Geldanlageverhalten für ein Panel geeignet ist. (3 Punkte)

Lösung:

Nicht geeignet: Werbeerinnerung, wird durch Abfrage verändert.

Geeignet: Einkaufsverhalten wird durch Abfrage nur wenig verändert.

Anlageverhalten: Geeignet, da es nur wenig durch Abfrage verändert wird.

c) Welches dieser Angebote ist vorzuziehen, wenn Sie...

c1) ...möglichst genau das Anlageverhalten des jeweils letzten Jahres messen wollen? (2 Punkte)

Lösung:

Wellenbefragung hat größere Stichprobe und misst daher Zustand genauer

c2) ...möglichst genau die **Veränderung** des Anlageverhaltens im letzten Jahr gegenüber dem Vorjahr messen wollen? Rechnerische Begründung! (5 Punkte)

Lösung:

$$10.000 / (1 - 0,76 * 0,68) = 20.695$$

Die Stichprobe der Wellenbefragung ist kleiner -> Panel misst Veränderung genauer

d) Das Institut, das die Wellenbefragung anbietet, teilt auf Nachfrage mit, dass die effektive Fallzahl bei 10.500 liegt.

d1) Wie ist die effektive Fallzahl definiert und wodurch entsteht die Abweichung der effektiven Fallzahl von der tatsächlichen Fallzahl? (4 Punkte)

Lösung:

Die effektive Fallzahl ist die Stichprobengröße einer optimalen Stichprobe, welche die gleiche Streuung liefert wie die vorliegende Stichprobe. Sie entstehen durch notwendige Gewichtungen, weil es z.B. bestimmte Stichprobenausfälle gibt und daher die Stichprobe verzerrt ist.

d2) Wieviel Prozent ungenauer ist die vorliegende Stichprobe gegenüber einer optimalen Stichprobe von 18.000 Haushalten? (2 Punkte)

Lösung:

$$\text{Wurzel}(10.500/18.000) = 0,764 \rightarrow 23,6\% \text{ ungenauer.}$$