

**Aufgabenblatt zu  
Statistische Methoden der soziologischen Bildungsforschung**

1. Geben Sie drei Unterschiede zwischen Fallstudien und statistischen Untersuchungen an.
2. Was versteht man unter ‘units of analysis’? Geben Sie drei verschiedene Beispiele an.
3. Welche beiden Ziele statistischer Methoden werden im Skript angegeben?
4. Was wird vorausgesetzt, wenn man von ‘statistischen Beschreibungen’ spricht?
5. Wie wird im Skript der Begriff ‘learning frame’ verwendet? Geben Sie drei Beispiele an.
6. Erläutern Sie anhand eines Beispiels den Diversity Index.
7. Es gibt 6 Studiengänge, in denen 10, 20, 50, 80, 70 bzw. 20 Personen eingeschrieben sind.
  - a) Berechnen und interpretieren Sie den Diversitätsindex.
  - b) Wie verändert sich der Diversitätsindex, wenn die ersten beiden Studiengänge zusammengefasst werden?
8. Es gibt 6 Studiengänge, in denen 10, 20, 50, 80, 70, 20 männliche und 30, 20, 30, 70, 80, 20 weibliche Studierende eingeschrieben sind. Berechnen und interpretieren Sie den Dissimilaritätsindex.
9. Es gibt 6 Studiengänge, in denen 10, 20, 50, 80, 70, 20 männliche und 45, 30, 45, 105, 120, 30 weibliche Studierende eingeschrieben sind. Berechnen Sie den Dissimilaritätsindex und vergleichen Sie das Ergebnis mit der vorangegangenen Aufgabe.
10. Welchen minimalen und welchen maximalen Wert kann der Diversitätsindex bei 10 Kategorien annehmen? Interpretieren Sie die beiden Werte.
11. Welchen minimalen und welchen maximalen Wert kann der Dissimilaritätsindex annehmen? Interpretieren Sie die beiden Werte.
12. Übersetzen Sie den Begriff ‘opportunity’ ins Deutsche. Welche Schwierigkeiten sind mit diesem Begriff verbunden? Diskutieren Sie anhand

eines Beispiels, wie man Opportunities feststellen kann.

13. Was versteht man unter ‘educational attainment’?
14. Wie wird im Skript der Begriff ‘deskriptive Generalisierung’ verwendet?
15. Erklären Sie anhand von Beispielen die Unterscheidung zwischen ‘data generating process’ und ‘fact generating process’.
16. Geben Sie drei verschiedene Beispiele für Regeln an.
17. Geben Sie jeweils ein Beispiel für eine statische und eine dynamische prediktive Regel an.
18. Erläutern Sie anhand eines Beispiels, wie man eine Regel sowohl verbal als auch mit Hilfe von Variablen formulieren kann.
19. Stellen Sie den Verlauf der für das Logitmodell verwendeten Funktion

$$v \longrightarrow \frac{\exp(v)}{1 + \exp(v)}$$

graphisch dar.

20. Für das Modell 2.3.1 werden folgende Daten angenommen:

$X$	$Z$	$Y=0$	$Y=1$	
0	1	300	400	
0	2	80	320	(1)
1	1	40	260	
1	2	80	720	

- a) Tabellieren Sie die Funktion

$$(x, z) \longrightarrow \Pr(\dot{Y} = 1 | \ddot{X} = x, \dot{Z} = z) \quad (2)$$

- b) Nehmen Sie ein saturiertes Modell an und berechnen Sie eine Varianzzerlegung. Wieviel Prozent der Varianz werden durch das Modell im statistischen Sinn erklärt?
21. Berechnen Sie mit Hilfe eines Statistikprogramms auf der Grundlage der Daten (1) für die Funktion (2) Logitmodelle:
  - a) Ein Logitmodell ohne einen Interaktionsterm. Berechnen Sie mit den geschätzten Modellparametern Werte der Funktion (2).

- b) Ein Logitmodell mit einem Interaktionsterm.  
Berechnen Sie mit den geschätzten Modellparametern Werte der Funktion (2).

22. Berechnen Sie mit den Daten der Table 3.1 ein Logitmodell ohne einen Interaktionsterm. Berechnen Sie mit dem geschätzten Modell die in (3.11) dargestellten Effekte. Erläutern Sie die Differenzen.
23. Erläutern Sie am Beispiel der vorangegangenen Aufgabe, in welchem Sinn es bei einem Logitmodell immer einen Interaktionseffekt gibt. Geben Sie den Interaktionseffekt zahlenmäßig an.
24. Erörtern Sie anhand des Modells 2.3.1 die Unterscheidung zwischen einem direkten und einem indirekten Effekt von  $\ddot{X}$  auf  $\dot{Y}$ . Wie können diese Effekte mit den in (1) angegebenen Daten berechnet werden?
25. Erläutern Sie anhand eines Beispiels den Begriff ‘Mediatorvariable’.
26. Erläutern Sie anhand eines Beispiels den Begriff ‘Moderatorvariable’.
27. Erläutern Sie anhand eines Beispiels den Begriff ‘Confounder’.
28. Beziehen Sie sich auf das Modell 5.1 und berechnen Sie mit den Daten in Table 5.1 die konditionalen und nicht-konditionalen Effekte von  $\ddot{Z}$  auf  $\dot{Y}_2$ . Vergleichen Sie diese Effekte mit den entsprechenden Effekten von  $\ddot{X}$ .
29. Leiten Sie aus dem Modell 5.1 ein reduziertes Modell ab, das die Variable  $\dot{Y}_1$  nicht mehr enthält. Berechnen Sie mit den Daten der Table 5.1 Werte für die Funktion

$$(x, z) \longrightarrow \Pr(\dot{Y}_2 = 1 | \ddot{X} = x, \ddot{Z} = z)$$

die dem reduzierten Modell entspricht.

30. Wie ist der Effekt von  $\ddot{X}$  auf  $\dot{Y}_2$  im reduzierten Modell (der vorangegangenen Aufgabe) zu definieren und zu interpretieren? Muss dabei  $\ddot{Z}$  berücksichtigt werden?
31. In Formel (5.25) wird der Wertebereich der vierdimensionalen Variablen als  $\{0, 1\}^4$  dargestellt. Geben Sie diese Menge explizit an.
32. Fassen Sie kurz zusammen, wodurch sich die beiden in Abschnitt 6.1 unterschiedenen Konzeptionen der Definition primärer und sekundärer Effekte unterscheiden.

33. Nehmen Sie an Stelle von Table 7.1.1 die folgenden Daten an:

		number of cases with		
$x$	$z$	$\dot{Y} = 1$	$\dot{Y} = 2$	$\dot{Y} = 3$
0	0	2000	1200	500
0	1	1600	900	700
1	0	600	400	900
1	1	300	100	800

- a) Berechnen Sie analog zu (7.4) die Wahrscheinlichkeiten.
- b) Berechnen Sie mit einem Statistikprogramm die Parameter des in (7.5) angegebenen multinomialen Logitmodells.
- c) Berechnen Sie auch die Parameter eines nicht-saturierten Logitmodells, das den Interaktionsterm nicht enthält.
- d) Berechnen Sie mit den in (b) und (c) geschätzten Parametern die Werte der in (a) tabellierten Wahrscheinlichkeiten.
34. Berechnen Sie mit den Parametern in (7.15) Schätzwerte für die in (7.14) definierten Wahrscheinlichkeiten, und zwar für  $\dot{Y} = 1, 2, 3, 4$  und  $\ddot{X} = 1, 2, 3$  und  $\ddot{Z} = 5$ . Stellen Sie die Werte übersichtlich in einer Tabelle mit 12 Feldern dar und geben Sie Interpretationen an.