

Tutorien

Statistik II

Serdar Selova, Alexander Lechner
T_EX: B. Philipp Kleer
Team Professur für Methoden der Politikwissenschaft

8. April 2022
Version: v2.1

Dieses Material ist ausschließlich für Studierende im BA Social Sciences an der Justus-Liebig-Universität Gießen konzipiert. Es ist als begleitendes Lernmaterial zur Vorlesung *Statistik II* konzipiert.

Hinweis: Da das *typesetting* der Formelsammlung mit englischem Sprachformat erfolgt, ist das Dezimalzeichen ein Punkt (.) und nicht ein Komma (wie im Deutschen üblich).

Inhaltsverzeichnis

1	Tutorium 1	2
2	Tutorium 2	4
3	Tutorium 3	5
4	Tutorium 4	6
5	Tutorium 5	7
6	Tutorium 6	8

1 Tutorium 1

Zur Erinnerung hier nochmal die Angabe zwischen Grundgesamtheit und Stichprobe:

	Grundgesamtheit	Stichprobe
Arithm. Mittel	μ	\bar{x}
Varianz	σ^2	s^2
Standardabweichung	σ	s

1. Was besagt das zentrale Grenzwerttheorem?
2. Welche Bedingungen müssen erfüllt sein, damit man von normalverteilten Stichprobenkennwerten ausgehen kann?
3. Welchen Zweck erfüllt die z-Standardisierung?
4. Beschreibe wie das Vorzeichen und der numerische Wert eines z-Wertes zu interpretieren ist.
5. Lisa hat in ihrer Statistikklausur 8 Punkte geschrieben. Der Mittelwert lag bei 6.5 und die Standardabweichung bei 3.15 Punkten. Treffe eine Aussage über Lisas relative Position innerhalb der Notenverteilung
6. Björn und Malte haben an verschiedenen Hotdog-Wettessen teilgenommen und wollen nun wissen, wer im Vergleich zu seinen Konkurrenten besser abgeschnitten hat. Malte aß in seinem Wettkampf 15 Hotdogs, während das arithmetische Mittel, bei einer Standardabweichung von 3 Hotdogs, nur 14 schaffte. Björn gelang es 17 Hotdogs zu essen, wobei bei seinem Wettkampf ein gleiches arithmetische Mittel von 14 Hotdogs erreicht wurde, die Standardabweichung lag aber bei 7 Hotdogs.
7. Welche Eigenschaften hat eine Standardnormalverteilung?
8. Wodurch zeichnet sich eine Zufallsstichprobe aus?
9. Erkläre in eigenen Worten wie du die z-Tabelle für einen Wert von $z = -0.05$ lesen musst.
10. Nimm an, du hast die Verteilung einer stetigen Variable vorliegen (z.B. Alter; $\mu = 45, \sigma = 13$). Wie hoch ist die Wahrscheinlichkeit, dass eine Person genau 70 Jahre alt ist $P(X = 70)$?
11. Bestimme den Flächenanteil einer Normalverteilung, der den folgenden Bereichen entspricht.
 - (a) $z \leq 0.25$
 - (b) $z > 0.8$
 - (c) $z \leq -1.50$
 - (d) $z \leq 0.91$
 - (e) $z > 1.27$
 - (f) $z \leq 0.69$
12. Finde für eine Normalverteilung denjenigen z-Wert, der die Verteilung (annähernd) in die folgenden Bereiche unterteilt:
 - (a) obere 20 %
 - (b) obere 60 %
 - (c) mittlere 70 %
 - (d) obere 10 %

- (e) obere 70 %
 - (f) mittlere 40 %
13. Gegeben sei eine Verteilung von IQ-Werten mit $\mu = 100$ und $\sigma = 15$. Wie groß ist die Wahrscheinlichkeit, zufällig eine Person mit einem $IQ \leq 120$ auszuwählen, wenn diese Verteilung normalverteilt ist?
14. Bestimme für eine Normalverteilung mit $\mu = 60$ und $\sigma = 12$ die nachfolgenden Wahrscheinlichkeiten.
- (a) $P(X > 66)$
 - (b) $P(X \leq 75)$
 - (c) $P(48 \leq X \leq 72)$
15. Angenommen die Werte eines standardisierten Leistungstests für Studierende seien normalverteilt mit dem Mittelwert $\mu = 70$ und der Standardabweichung $\sigma = 10$. Wie groß ist der Anteil Studierender mit einem Wert $X \leq 75$ ($P(X \leq 75)$)?
16. Gegeben seien normalverteilte Populationswerte eines IQ-Tests mit einem Mittelwert $\mu = 100$ und einer Standardabweichung von $\sigma = 15$. Wie groß ist die Wahrscheinlichkeit bei einem Stichprobenumfang von $n = 9$ einen Mittelwert von $\bar{x} > 97$ zu erhalten?
17. Wie hoch ist bei einer linksschiefen Verteilung mit $\mu = 60$ und $\sigma = 12$ die Wahrscheinlichkeit, einen Wert größer 45 zu erhalten?

2 Tutorium 2

1. Definiere den Standardfehler!
2. Welche Funktion erfüllt der Standardfehler?
3. Kreuze die richtige Antwort an! Der Standardfehler verringert sich, je ...
 - (a) ... größer der Mittelwert in der Population.
 - (b) ... größer der Mittelwert in der Stichprobe.
 - (c) ... kleiner die Varianz in der Stichprobe.
 - (d) ... größer die Stichprobe.
4. Welche Eigenschaften hat die t-Verteilung zu?
5. Wann nutzt man die t-Verteilung?
6. Gegeben seien Daten mit einer Stichprobengröße von $n = 7$, einem Mittelwert von $\mu = 8$ und einer Standardabweichung von $\sigma = 4.2$. Welchen Wert nimmt der Standardfehler an?
 - (a) 2.0507
 - (b) 2.6527
 - (c) 5.8123
 - (d) 1.5875
7. Mittels einer Zufallsstichprobe ($n = 525$) wurde für ein Land eine mittlere Links-Rechts-Einschätzung von $\bar{x} = 4.78$ ermittelt ($s_x = 1.75$) (0: *sehr links*, 10: *sehr rechts*). Die Standardabweichung der Population beträgt $\sigma_x = 1.42$). Berechne das Konfidenzintervall, in dem der wahre Mittelwert mit einer 90 % Wahrscheinlichkeit liegt mithilfe der z-Verteilung.
8. Mittels einer Zufallsstichprobe ($n = 101$) wurde für ein weiteres Land eine mittlere Zufriedenheit mit der Demokratie von $\bar{x} = 6.78$ ermittelt ($s_x = 2.43$) (0: *gar nicht zufrieden*, 10: *sehr zufrieden*). Berechne das Konfidenzintervall, in dem der wahre Mittelwert mit einer 95 % Wahrscheinlichkeit liegt mithilfe der t-Verteilung.

3 Tutorium 3

1. Die Universität will überprüfen, ob Studierende, die an Statistiktutorien teilnehmen, im Mittel andere Noten in der Abschlussklausur erzielen, als Studierende, die nicht an den entsprechenden Angeboten teilnehmen. Formuliere die H_0 und H_A .
2. Was ist ein p-Wert?
3. Du hast zwei statistische Tests durchgeführt. Das Signifikanzniveau beträgt jeweils $\alpha = 0.05$. Entscheide in den zwei Szenarien hinsichtlich der Nullhypothese (H_0).
 - (a) p-Wert < 0.05
 - (b) p-Wert > 0.05
4. Durchschnittlich waren Arbeitslose in Ostdeutschland im ALLBUS 1998 ($n = 144$) seit 67.7 Wochen arbeitslos. Die Varianz der Grundgesamtheit beträgt $\sigma_x^2 = 61.5$. Berechne ein 95-prozentiges Konfidenzintervall und interpretiere dieses.
5. Welche Fehlschlüsse sind bei Hypothesentest möglich? Erläutere diese!
6. Einflussfaktoren auf Hypothesentests sind:
 - (a) Streuung und Konfidenzintervall
 - (b) Stichprobengröße und Streuung
 - (c) Stichprobengröße und Konfidenzintervall
7. Welche Aussagen sind richtig? Diese Faktoren haben einen Einfluss auf Hypothesentest, weil ...
 - (a) ... das Konfidenzintervall ein Intervall von der Grundgesamtheit aufgrund einer Stichprobe angibt.
 - (b) ... die Streuung den Standardfehler beeinflusst.
 - (c) ... die Stichprobengröße die Höhe des Standardfehlers mitbestimmt.

4 Tutorium 4

1. Ein Bekannter von dir stellt die Behauptung auf, dass Arbeitslose in Ostdeutschland durchschnittlich ein Jahr (52 Wochen) arbeitslos seien. Aufgrund der Arbeitsmarktsituation in Ostdeutschland vermutest du jedoch, dass die durchschnittliche Dauer der Arbeitslosigkeit nicht bei einem Jahr liegt. Deine Behauptung stellt die Alternativhypothese dar. Im ALLBUS 1998 wurde die Dauer der Arbeitslosigkeit erfasst. Durchschnittlich waren die 144 ostdeutschen Befragten, die Arbeitslosigkeit angaben, seit 67.7 Wochen arbeitslos. Die Varianz der Grundgesamtheit beträgt $\sigma_x^2 = 61.5$. Prüfe auf Basis der ALLBUS Daten auf einem Signifikanzniveau von 5 %, ob die durchschnittliche Dauer der Arbeitslosigkeit nicht ein Jahr, also 52 Wochen, beträgt.
2. Die Universität will überprüfen, ob die Einführung eines Tutoriums die Abschlussnoten im Mittel verbessert. Das Ergebnis einer Stichprobe ($n = 40$) von Studierenden, die am Tutorium teilgenommen haben, zeigt, dass der Mittelwert $\bar{x} = 9.4$ Punkte beträgt. Der Mittelwert in der Grundgesamtheit liegt bei $\mu_x = 8.5$ Punkten und hat eine Standardabweichung von $\sigma_x = 5$. Führe einen gerichteten Hypothesentest mit einer Signifikanz von $\alpha = 0.05$ durch.
3. Wofür werden Tests für Mittelwertunterschiede zwischen einer Stichprobe und einer Grundgesamtheit genutzt? Überlege dir Beispiele.

5 Tutorium 5

1. Was sind Freiheitsgrade (df, v)?
2. Wie leitet man Freiheitsgrade her?
3. Warum werden Freiheitsgrade (df, v) statt n genutzt?
4. Du bist in der Bildungsforschung und bekommst den Auftrag an deiner Universität die Ergebnisse einer Statistiklausur von Studierenden des Fachbereichs Volkswirtschaft und Soziologie auf Unterschiede zu untersuchen. Dabei sollst du mit einem Signifikanzniveau von $\alpha = 0.05$ eine Aussage treffen. Die Stichprobe des Fachbereichs Soziologie umfasst $n_1 = 214$ Studierende und die des Fachbereichs Volkswirtschaft $n_2 = 288$ Studierende. Die Soziologiestudierenden haben durchschnittlich 8.4 ($\bar{x}_1 = 8.4$) geschrieben, während die VWL-Studierenden im Mittelwert $\bar{x}_2 = 6.3$ Punkte erreicht haben. Die Varianz lag in der Soziologie wie in der VWL bei $s_1^2 = s_2^2 = 25$. Führe einen Mittelwertvergleich durch! Überlege dir zuerst, welche Testsituation vorliegt!

6 Tutorium 6

1. Wodurch unterscheiden sich abhängige und unabhängige Stichproben?
2. Es soll überprüft werden, ob sich die Mittelwerte eines Merkmals, das an zwei abhängigen Stichproben gemessen wurde, signifikant unterscheiden. Es wird davon ausgegangen, dass es zwischen dem ersten und dem zweiten Messzeitpunkt zu einem Anstieg in dem interessierenden Merkmal kommt. Die Stichprobe hat einen Umfang von $n = 251$ je Messzeitpunkt. Das arithmetische Mittel der Mittelwertdifferenzen zwischen den beiden Messzeitpunkten beträgt $\bar{x}_d = 11$; die Standardabweichung wird mit $\hat{\sigma}_d = 2$ angegeben. Das Signifikanzniveau ist $\alpha = 0.01$. Der Test soll einseitig getestet werden. Führe den passenden Test durch!