

Statistische Inferenz | 02 | Zentraltendenz, Streuung, Standardfehler

Prof. Dr. Roland Schäfer | Germanistische Linguistik FSU Jena

30. Oktober 2024

1 Skalenniveaus

Bestimmen Sie das Skalenniveau von folgenden Messgrößen:

1. Prozentwerte
2. Wortfrequenz-Rang (häufigstes Wort, ..., seltenstes Wort)
3. Kasus
4. Geschwindigkeit
5. Akzentsitz (Erstsilbe, Mittelsilbe, Endsilbe)
6. Satzlänge, gemessen in Wörtern
7. Frequenz eines Wortes im Korpus (absolute Zahl)
8. Höhe über NN
9. DSH-Prüfungsniveau (I – III)
10. Verhältnis Satzlänge in Wörtern zu Wortlänge in Silben in einem Text
11. Wortklasse (= Wortart)
12. Beschleunigung
13. Textniveau (leicht, mittel, schwer)
14. Frequenz eines Wortes im Korpus pro eine Millionen Wörter
15. Textsorte

2 Modus und Median

Ermitteln von Hand (ohne R!) Sie den Modus und wo möglich den Median für folgende Messreihen:

1. $x = [\text{Nom, Akk, Akk, Akk, Nom, Dat, Gen, Nom, Nom, Akk, Dat, Dat, Akk, Akk}]$
2. $x = [4, 5, 3, 3, 3, 2, 1, 2, 2, 1, 5, 4, 2, 2, 1, 3, 2]$
3. $x = [4.3, 5.0, 3.0, 3.3, 3.7, 2.3, 1.3, 2.7, 2.0, 1.0, 5.0, 4.3, 2.0, 2.0, 1.3, 3.0, 2.7]$

3 Mittel und Streuung

Ermitteln Sie zuerst von Hand (gerne „von Hand in R“) für eine der Messreihen und dann in R für alle Messreihen das arithmetische Mittel, die Varianz und die Standardabweichung:

1. $x = [2.73, 1.85, 21.24, 17.97, 5.49, 18.90, 12.46, 0.97, 6.45, 7.43]$
2. $x = [1.00, 1.91, 3.12, 4.38, 4.72, 5.29, 3.82, 3.25, 2.04, 0.93]$
3. $x = [1.07, 1.06, 0.94, 1.84, 3.04, 3.22, 4.18, 5.27, 6.27, 6.75]$

4 z-Werte und Standardfehler

Ermitteln Sie für die Messreihen aus Aufgabe 3 die z-Werte für die Messpunkte und die Standardfehler (entweder von Hand oder „von Hand in R“). Formulieren Sie in eigenen Worten (jeweils ein Satz), was z-Werte und Standardfehler angeben.

5 Konfidenzintervalle (Anteilswerte)

Berechnen Sie für folgende Anteilswerte (alle p_i) die Konfidenzintervalle bei den Stichprobengrößen $n = 10$ und $n = 100$ auf den Konfidenzniveaus 0.9 und 0.99. Benutzen Sie zum Auffinden der kritischen Werte die Quantil-Funktion der Normalverteilung `qnorm()` in R.

1. $p = [0.21, 0.79]$
2. $p = [0.49, 0.51]$
3. $p = [0.12, 0.71, 0.17]$