

Zu den kumulierten Häufigkeiten bei gruppierten Merkmalen

Präziser sollte man schreiben:

Bei gruppierten, intervallskalierten Merkmalen tritt folgendes zusätzliche Problem auf: Die kumulierte Häufigkeitsverteilung hängt davon ab, wie die einzelnen Ausprägungen innerhalb der Intervalle verteilt sind.

Da diese unbekannt ist, erhält man nur untere und obere Schranken für die Werte der kumulierten Häufigkeitsverteilungen. Wie definiert man darauf aufbauend $H(x)$ und $F(x)$?

...

“Hausaufgabe:” Anwenden der Approximationsformel auf $H(40)$:

Geradengleichung:

$$\begin{aligned} H(x) &\approx \sum_{\ell=1}^{j-1} h_{\ell} &+& \frac{h_j}{(c_j - c_{j-1})} \cdot (x - c_{j-1}) \\ &\text{Ausgpkt.} && \text{Steigung} && \text{Wegstrecke} \\ \\ H(40) &\approx 53 &+& \frac{78}{48.5 - 35.5} \cdot (40 - 35.5) = \\ &= 80 && && \end{aligned}$$