

## Übung (2)

- (1) Nehmen Sie Ihre Geschwisterzahlverteilung aus Ihrer Übungsgruppe wieder vor. Fertigen Sie dazu eine geeignete Wertetabelle für die Verteilungsfunktion, und skizzieren Sie deren Graphen.
- (2) Hier sind 'gruppierte Daten' zur Verteilung der Rededauer der Teilnehmer von Diskussionsgruppen:

Dauer (Minuten)	0 – 1	1 – 3	3 – 5	5 – 10
relative Häufigkeit	0.55	0.2	0.15	0.1

- (a) Fertigen Sie eine Tabelle der zugehörigen Dichtewerte, und zeichnen Sie ein Histogramm dazu, beschriften Sie das richtig. Erscheint Ihnen die Form der Verteilung plausibel?
  - (b) Berechnen Sie näherungsweise (nicht genau, das ginge gar nicht, wieso?) den Mittelwert (d.h. genauer den Erwartungswert) und die Streuung. Erinnerung: Beide Parameter sind für die gemeinte grobe Rechnung so zu bestimmen, als lägen alle Werte jeweils in der Mitte des zugehörigen Gruppierungsintervalls.
  - (c) Welcher Populationsanteil liegt im Bereich  $\mu \pm 2\sigma$  in diesem Falle (mit den genäherten Verteilungsparametern aus )?
  - (d) Fertigen Sie eine geeignete Wertetabelle für die Verteilungsfunktion. Skizzieren Sie den Graphen der Verteilungsfunktion, und bestimmen Sie graphisch den Median der Verteilung. Vergleichen Sie mit dem Erwartungswert.
  - (e) Welche Steigung hat die Verteilungsfunktion im Intervall  $[1, 3]$ ? Wie hängt das mit der Dichtefunktion (im Sinne des Histogrammbildes idealisiert) zusammen?
- (3) Überlegen Sie eigene konkrete Beispiele für folgende Verteilungsformen: Glockenform, symmetrisch und schief, zweigipflige Verteilung (mit starkem Abfall zu den Rändern), J-Form beider Richtungen.
  - (4) Stellen Sie sich die verschiedenen typischen Verteilungsformen auf Intervallen derselben Breite vor (eine grobe Skizze ist jeweils günstig). Welche der Formen bewirken größere, welche kleinere Streuungen?
  - (5) **Hausaufgabe:**
    - (a) Stellen Sie die bisher betrachteten Möglichkeiten zusammen, eine Verteilung (von relativen Häufigkeiten oder Wahrscheinlichkeiten) zu definieren (bzw. zu beschreiben) durch eine Funktion (Zuordnung), und stellen Sie die graphischen Darstellungen davon daneben, wie sie typisch aussehen. Formulieren Sie dazu selber noch einmal in Worten, wie man die relativen Häufigkeiten bzw. Wahrscheinlichkeiten (**wofür jeweils??**) daraus abliest. Den Erfolg dieser Wiederholungsarbeit können Sie nunmehr daran absehen, wie leicht es Ihnen fällt, folgende Verständnisfragen zu beantworten:
    - (b)
      - (i) Warum kann eine Dichtefunktion in einem Intervall fallen, eine Verteilungsfunktion aber niemals?
      - (ii) Wenn man zu einer Variablen  $X$  die Verteilungsfunktion  $F_X$  kennt, wie kann man dann für beliebige Zahlen  $a < b$  den Wert von  $P(a < X \leq b)$  ermitteln? Zusatzfrage: In welchen Fällen ergibt sich für die Frage nach  $P(a \leq X \leq b)$  (wieder mit Kenntnis von  $F_X$ ) eine Schwierigkeit, in welchen nicht?
      - (iii) Eine Variable  $X$  nehme nur die Werte  $a_1 < a_2 < \dots < a_n$  an. Ferner kenne man alle Zahlen  $p_i = P(X = a_i)$  für  $1 \leq i \leq n$ . Wie kann man dann  $F_X(a)$  ausrechnen für  $a_i < a < a_{i+1}$ ? (dabei sei  $i$  irgendeine natürliche Zahl im Bereich von 1 bis  $n - 1$ .) Zusatzfrage: Was ist  $F_X(a_n)$  nach Definition?
      - (iv) Sie haben eine Variable  $X$ , die mit einer Dichte verteilt ist, also ein Kontinuum von Werten annimmt, und bei der stets die auf ein Intervall entfallende relative Häufigkeit (oder Wahrscheinlichkeit) durch den Flächeninhalt unter dem Dichtegraphen im betreffenden Intervall gegeben ist. Warum nützt die Information  $P(X = a)$  nicht das Geringste? Was bedeutet es für die Streuung, wenn die Verteilungsfunktion innerhalb eines sehr schmalen Intervalls von fast Null auf fast Eins ansteigt (also sehr steil)? Wie verhält sich die Verteilungsfunktion in einem Bereich sehr niedriger Dichtewerte?