Lösungen zu Aufgaben 3

Aufgabe 1)

a) *Bitte erstellt eine Häufigkeitstabelle, in der ihr die relativen Häufigkeiten [p_k] sowie die kumulierten prozentualen Anteile [$cp_{\%}(x_k)$] berechnet!*

X	n_k	p_k	$cp_{\%}(x_k)$
0	2	0.2	20%
1	3	0.3	50%
2	3	0.3	80%
3	2	0.2	100%

b) *Wie viel Prozent der Teilnehmer haben höchstens zwei mal ein Buch zur Hand genommen?*

80% (ablesbar aus Tabelle → $cp_{\%}(x_{k=2})$)

c) *Wie viel Prozent der Teilnehmer haben mindestens zwei mal ein Buch zur Hand genommen?*

50% (100% - 50% ($cp_{\%}(x_{k=1})$ → die kumulierten prozentualen Anteile der letzten nicht mit eingeschlossenen Ausprägung

oder: 0.3($p_{k=2}$) + 0.2($p_{k=3}$))

Aufgabe 2)

a) *Definitionen*

Modus: Der Modus einer Verteilung ist diejenige Ausprägung, die am häufigsten vorkommt (Notation: x_{mod}). Bei klassierten Daten spricht man von Modalklassen.

Median: Der Median einer Verteilung ist diejenige Ausprägung, die in der Mitte der nach ihrer Größe sortierten Messwerte steht. Der Median teilt die Verteilung einer Variablen exakt in zwei Hälften (Notation: $\tilde{x}$).

Arithmetisches Mittel: Das arithmetische Mittel wird berechnet als Summe aller Realisationen (Messwerte) einer Verteilung geteilt durch die Anzahl der Realisationen (Notation: $\overline{X}$).

b) *Unter welchen Überbegriff lassen sich die Termini einordnen?*

Maße der zentralen Tendenz (oder auch: Lagemaße)

"Durch welchen Wert wird die Merkmalsverteilung am besten repräsentiert?"

Der Begriff "Tendenz" legt nahe, dass man bei der Interpretation der verschiedenen Mittelwerte die inhaltliche Aussagekraft nicht überschätzen sollte. Mittelwerte sind ohne die Einbeziehung von Streuungsmaßen für sich alleine genommen oft wenig aussagekräftig (siehe auch Beispiel in der VL [→ ‚Sexualpartner'] bzgl. der möglichen Differenz zwischen den verschiedenen Mittelwerten).

c) *Welche Skalenniveaus müssen theoretisch gesehen für die drei verschiedenen Mittelwerte jeweils mindestens vorausgesetzt sein?*

Modus: nominales Skalenniveau

Median: ordinales Skalenniveau

Arithmetisches Mittel: metrisches Skalenniveau (in der Praxis werden ordinal-skalierte Variablen häufig als metrisch behandelt; daher wird häufig auch bei ordinalen Skalenniveaus das arithmetische Mittel berechnet)

Aufgabe 3)

a) *Bestimmt oder berechnet für die Häufigkeitstabelle aus ‚Aufgabe 1' bitte alle sinnvollen (bzw. theoriegeleitet „erlaubten") Mittelwerte!*

Modi: $X_{mod1}=1$; $X_{mod2}=2$
(kann abgelesen werden, es ist also keine Berechnung im eigentlichen Sinne notwendig)

Median: $\tilde{x}=1{,}5$

$$\frac{1}{2}\left(X\left(\frac{10}{2}\right)+X\left(\frac{10}{2}+1\right)\right)$$
$$=\frac{1}{2}(X(5)+X(6))$$

relevant sind demnach die Merkmalsausprägungen des 5. Und 6. Falls, also:

$$\frac{1}{2}(1+2)=1{,}5$$

Arithmetisches Mittel: $\bar{x}$ **=1,5**

$$\frac{1}{10}(0+3+1+1+2+0+3+1+2+2)=1{,}5$$
$$oder:$$
$$\frac{1}{10}((2\times 0)+(3\times 1)+(3\times 2)+(2\times 3))=1{,}5$$

b) *Für den Fall, dass ihr kein arithmetisches Mittel berechnet habt, holt dies bitte nach! Die Summe aller Abweichungen der Einzelwerte muss Null ergeben:*

$$\sum_{i=1}^{n}\left(x_i-\bar{x}\right)=0$$

Bitte überprüft euer Ergebnis rechnerisch anhand der Formel!

$$(0-1{,}5)+(3-1{,}5)+(1-1{,}5)+(1-1{,}5)+(2-1{,}5)+(0-1{,}5)+(3-1{,}5)+(1-1{,}5)+(2-1{,}5)+(2-1{,}5)$$
$$=-1{,}5+1{,}5-0{,}5-0{,}5+0{,}5-1{,}5+1{,}5-0{,}5+0{,}5+0{,}5$$
$$=0$$

c) *Bestimmt oder berechnet für die nachfolgende grafische Darstellung alle sinnvollen (bzw. theoriegeleitet „erlaubten") Mittelwerte!*

Modus: $X_{mod}=2$
(kann problemlos aus der Grafik abgelesen werden)

Median: $\tilde{x}=2$
Mögliche Vorgehensweise:
Häufigkeitstabelle erstellen.

X	n_k	p_k	$cp_{\%}(x_k)$
1	181	0.147	14,7%
2	475	0.386	53,3%
3	332	0.270	80,3%
4	195	0.158	96,1%
5	48	0.039	100,0%

n ist ungerade. Daher kann der Median problemlos abgelesen werden. Bei den kumulierten prozentualen Anteilen wird die 50%-Marke bei X=2 überschritten.

Elaborierter:

$$X(\frac{1231+1}{2}) = X(616)$$

Der Median liegt also bei der 616ten Realisation (Fall) der geordneten Verteilung.
181(X=1)+475(X=2) ergibt bereits 656. Also liegt der Median bei X=2.

<u>Arithmetisches Mittel</u>: Streng genommen handelt es sich hier um eine ordinal-skalierte Variable. Die Abstände der Merkmalsausprägungen sind nicht interpretierbar, daher wäre die Berechnung eines arithmetischen Mittels theoriegeleitet als nicht sinnvoll anzusehen. In der Praxis würde es für eine solche Skala allerdings häufig errechnet:

$\overline{X}$=2,556

$$\frac{1}{1231}(181 \times 1 + 475 \times 2 + 332 \times 3 + 195 \times 4 + 48 \times 5)$$
$$= 0{,}000812348(181 + 950 + 996 + 780 + 240)$$
$$= 0{,}000812348 \times 3147$$
$$= 2{,}556$$