Frage:

Im Arbeitsbuch (2te Auflage), Seite 17, Aufgabe 2.4, Teilaufgabe 5 mit folgenden Tabellenwerte (siehe Häufigkeitstabelle), wird verlangt $H(1 <= x < 5)$ aufzurechnen und laut Formel ist das für diskrete und metrisch skalierte Merkmale mit der Formel $F(5) - F(1) + f(1) - f(5)$ zu berechnen! In den Lösungen wird nur $F(4) - F(0)$ gerechnet. Wie kommt man auf diese einfachere Lösung?

Antwort:

In dieser Aufgabe geht es um das Merkmal: X = Anzahl der geschossenen Tore während der WM pro Spiel.
Nachfolgend die Häufigkeitstabelle und die Grafik der empirischen Verteilungsfunktion des Merkmals X:

Anzahl der Tore	absolute Hkeit a_j	relative Hkeit $f_j(x)$	emp Vtlgsfkt$F_j(x)$
0	7	10,9	10,9
1	13	20,3	31,3
2	18	28,1	59,4
3	12	18,8	78,1
4	10	15,6	93,8
5	2	3,1	96,9
6	2	3,1	100
Gesamt	64	100	

Tabelle 1: Häufigkeitstabelle der geschossenen Tore

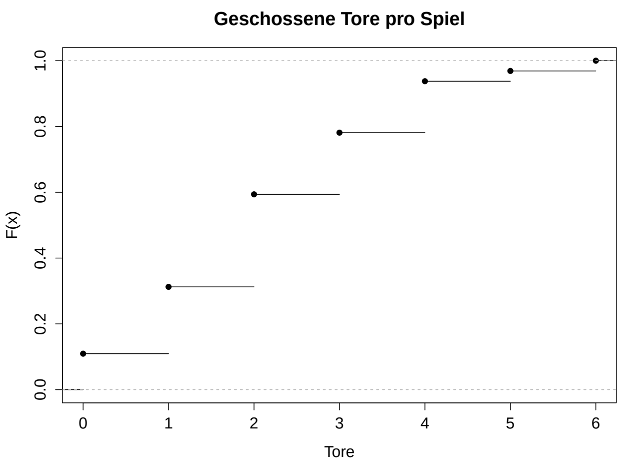

Abbildung 1: Empirische Verteilungsfunktion

Gesucht ist bei dieser Teilaufgabe die Anteil der Spiele der WM bei welchen mindestens 1 Tor, aber weniger als 5 Tore geschossen wurden.

Also,

$$H(1 \leq X < 5), \tag{1}$$

dies lässt durch zwei Möglichkeiten berechnen die auf inhaltlichen Überlegungen basieren:

1. Möglichkeit:

$$H(1 \leq X < 5) = F(5) - F(1) + f(1) - f(5) = 0{,}969 - 0{,}313 + 0{,}203 - 0{,}031 = 0{,}828 \tag{2}$$

Also bei ca 83% der Spiele fielen mindestens 1, jedoch weniger als 5 Tore.

Inhaltliche Interpretation der Berechnung:

Zuerst wird der Anteil der Spiele bei welchen bis zu 5 Tore (0,1,2,3,4,5) geschossen wurden herangezogen, $F(5) = 96{,}9\%$ aller Spiele.
Jetzt werden alle Spiele abgezogen bei denen höchstens 1 Tor geschossen wurde $F(1) = 31{,}3\%$ aller Partien. Das sind alle torlosen Spiele, sowie Spiele mit nur 1 Tor.
Hierduch ergibt sich nun der Teil der Spiele bei welchen 2,3,4,5 Tore geschossen wurden.
Um die interessierende Menge der Spiele mit 1,2,3,4 Toren zu erhalten muss der Anteil der Spiele mit einem Tor addiert werden $f(1) = 20{,}3\%$. Die Spiele bei welchen 5 Tore geschossen wurden,$f(5) = 3{,}1\%$, müssen abgezogen werden.

2. Möglichkeit:

$$H(1 \leq X < 5) = F(4) - F(0) = F(4) - f(0) = 93{,}8 - 10{,}9 = 0{,}829 \tag{3}$$

Also bei ca 83% der Spiele fielen mindestens 1, jedoch weniger als 5 Tore.

Inhaltliche Interpretation der Berechnung:

$F(4) = 93{,}8\%$, repräsentiert den Anteil aller Spiele bei denen höchstens 4 Tore gefallen sind, also 0,1,2,3,4.
Hiervon muss jetzt der Anteil aller Spiele abgezogen werden die torlos endeten, $f(0) = F(0) = 10{,}9\%$.

Fazit:

Beide Formeln sind für dieses Merkmal äquivalent. Man kann die gesuchte Häufigkeit auch folgendermaßen umformen:

$$H(1 \leq X < 5) = H(0 < X < 5) = H(0 < X \leq 4) = F(4) - F(0)$$

Damit landet man also auch so direkt bei der anderen, einfacheren Formel. Die nächst-kleinere Ausprägung nach 5 ist ja schließlich 4, damit ist < 5 gleichbedeutend mit ≤ 4 und es gilt $F(4) + f(5) = F(5)$. Ebenso gilt $F(0) + f(1) = F(1)$.

Frage:

Außerdem ist nach der längeren Formel das Ergebnis 0.828 und bei der vereinfachten Formel 0.829? Werden solche Unterschiede im Ergebnis je nach Rechenformel unterschieden? Bei MC Aufgaben wäre somit die Angabe der richtigen Antwort ja falsch

Antwort:

Dieser Unterschied basiert auf dem „Rundungsfehler“ in der Häufigkeitstabelle, bei Verwendung der ungerundeten relativen Häufigkeiten, bzw. kumulierten relativen Häufigkeiten würde sich dieser Fehler nicht ergeben und es würde exakt dasselbe herauskommen. Bei MC Aufgaben wird natürlich darauf möglichst darauf geachtet dass es bei den Antwortmöglichkeiten nur eindeutig richtige bzw falsche Antwortmöglichkeiten gibt.

Frage:

Auf wie viele Nachkommastellen muss in der Klausur gerundet werden?

Antwort:

Die Anzahl der Nachkommastellen auf die gerundet werden muss beträgt 2, das auch steht auf dem Deckblatt der Klausur, vgl. hierzu auch die Probeklausur auf der Homepage zur Veranstaltung.