Statistik I
WS 2013/2014
Musterlösung 4
Aufgaben 12-15

Aufgabe 12

(a) Man geht von der nachstehenden Arbeitstabelle aus:

Klasse	k	Klassenobergrenze x_k^*	Klassenmitte $\frac{x_k^*+x_{k-1}^*}{2}$	absolute Häufigkeiten n_k
	0	1		
1-5	1	5	3	22
6-8	2	8	7	34
9-11	3	11	10	45
12-14	4	14	13	52
15-19	5	19	17	38
20-30	6	30	25	29
31-45	7	45	38	20
				$\sum_{k=1}^{7} n_k = 240$

Betrachtet wird nun ausschließlich die Stichprobe der Raucher ($n_R = 240$). Als Näherung $\bar{x}'_R$ für den durchschnittlichen täglichen Zigarettenkonsum $\bar{x}_R$ der $n_R = 240$ Raucher errechnet man:

$$\begin{aligned}
\bar{x}'_R &= \frac{1}{n_R} \sum_{k=1}^{K} n_k \cdot \frac{x^*_k + x^*_{k-1}}{2} \\
&= \frac{1}{240}(22 \cdot 3 + 34 \cdot 7 + \ldots + 20 \cdot 38) \\
&= \frac{3561}{240} \\
&= 14,8375
\end{aligned}$$

(b) Der Stichprobenumfang wächst nun gegenüber der ersten Teilaufgabe.

$$\begin{aligned} n &= n_R + n_{NR} \\ &= 240 + 260 \\ &= 500 \end{aligned}$$

Der Zigarettenkonsum eines Nichtrauchers ist 0. Damit gilt auch für das arithmetische Mittel $\bar{x}_{NR}$ des täglichen Zigarettenkonsums für die $n_{NR} = 260$ Nichtraucher: $\bar{x}_{NR} = 0$.

Die beiden Teilgesamtheiten der Raucher und der Nichtraucher sind disjunkt, d.h. es gibt keine Überschneidungen, sodass man als Näherung $\bar{x}'$ für das arithmetische Mittel $\bar{x}$ des täglichen Zigarettenkonsums aller Befragten erhält:

$$\begin{aligned}
\bar{x}' &= \tfrac{1}{n}(n_R \cdot \bar{x}'_R + n_{NR} \cdot \bar{x}'_{NR}) \\
&= \tfrac{1}{n_R + n_{NR}}(n_R \cdot \bar{x}'_R + n_{NR} \cdot \bar{x}_{NR}) \\
&= \tfrac{1}{500}(240 \cdot 14{,}8375 + 260 \cdot 0) \\
&= 7{,}122
\end{aligned}$$

Aufgabe 13

(a) Bei $n = 400$ aufsteigend geordneten Urwerten $x_{[i]}$ $(i = 1, \ldots, n)$ ist der Median $\bar{x}^{\text{Med}}$ definiert als:

$$\begin{aligned}\bar{x}^{\text{Med}} &= \frac{1}{2}\left(x_{[\frac{n}{2}]} + x_{[\frac{n}{2}+1]}\right)\\ &= \frac{1}{2}\left(x_{[200]} + x_{[201]}\right)\end{aligned}$$

Bis zur Untergrenze der Einfallsklasse des Medians liegen laut Angabe $\widehat{F}(x^*_{k-1}) \cdot n = \widehat{F}(50) \cdot 400 = 0,480 \cdot 400 = 192$ Variablenwerte. Innerhalb der Einfallsklasse des Medians liegen entsprechend die Variablenwerte $x_{[193]}, \ldots, x_{[202]}$. Damit ergibt sich der Median $\bar{x}^{\text{Med}}$ unter Berücksichtigung der angegebenen Urwerte mit:

$$\begin{aligned}\bar{x}^{\text{Med}} &= \frac{1}{2}\left(x_{[200]} + x_{[201]}\right)\\ &= \frac{1}{2}(51,7 + 51,9)\\ &= 51,8\end{aligned}$$

(b) Die Anzahl der beobachteten Variablenwerte ist mit $n = 400$ so groß, dass $\widehat{F}(\bar{x}^{\text{Med}}) = 0,5$ gesetzt werden kann. Als Näherung $\bar{x}^{\text{Med}'}$ für den Median $\bar{x}^{\text{Med}}$ ermittelt man damit:

$$\begin{aligned}
\bar{x}^{\text{Med}'} &= x_{k-1}^* + \frac{0,5 - \widehat{F}(x_{k-1}^*)}{\widehat{F}(x_k^*) - \widehat{F}(x_{k-1}^*)} \cdot (x_k^* - x_{k-1}^*) \\
&= 50,0 + \frac{0,5 - 0,480}{0,505 - 0,480} \cdot (60,0 - 50,0) \\
&= 58,0
\end{aligned}$$

(c) Die $n_k = 10$ Merkmalswerte in der Einfallsklasse des Medians liegen alle sehr nahe an der Klassenuntergrenze. Sie sind also nicht gleichmäßig über die Klasse verteilt, was bei der Ermittlung der Näherung $\bar{x}^{\mathrm{Med}'}$ mittels linearer Interpolation angenommen wird. Der exakte Median aus den Urwerten und die Näherung aus der klassierten Häufigkeitsverteilung müssen hier also relativ weit auseinanderliegen.

Aufgabe 14

(a) Die mittlere jährliche Absatzsteigerung $\bar{w}^*$ erhält man aus dem durchschnittlichen Wachstumsfaktor $\bar{w}^{\text{Geo}} - 1$. $\bar{w}^{\text{Geo}}$ erhält man durch:

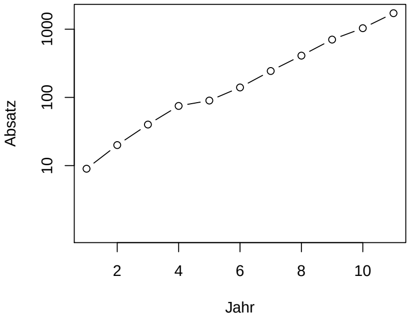
$$\begin{aligned}
\bar{w}^{\text{Geo}} &= \sqrt[10]{\frac{x_2}{x_1} \cdot \frac{x_3}{x_2} \cdot \frac{x_4}{x_3} \cdot \ldots \cdot \frac{x_{11}}{x_{10}}} \\
&= \sqrt[10]{\frac{x_{11}}{x_1}} \\
&= \sqrt[10]{\frac{1718}{9}} \\
&\approx 1{,}6907
\end{aligned}$$

Damit erhält man:

$$\begin{aligned}
\bar{w}^* &= \bar{w}^{\text{Geo}} - 1 \\
&= 69{,}07\%
\end{aligned}$$

(b) Für die Grafik verwendet man nachfolgende Tabelle.

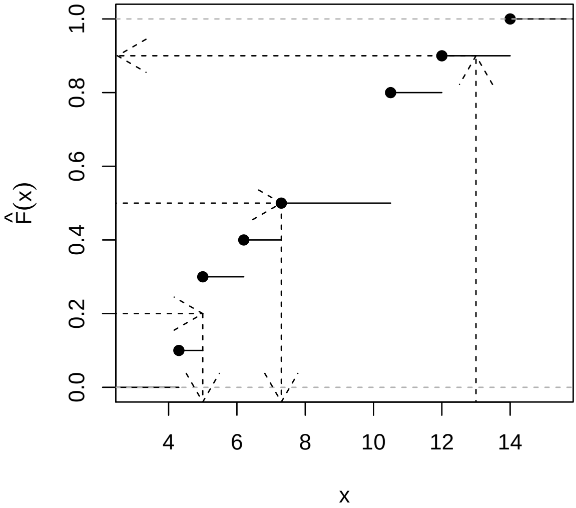
Jahr i	1	2	3	4	5	6	7	8	9	10	11
Absatz x_i	9	20	40	75	90	140	244	410	705	1036	1718
$\lg(x_i)$	0,95	1,30	1,60	1,88	1,95	2,15	2,39	2,61	2,85	3,02	3,24



Aufgabe 15

(a) Für die Grafik verwendet man nachfolgende Tabelle.

i	x_i	n_i	h_i	$H(x_i)$	$\widehat{F}(x_i)$
1	4,3	1	0,1	1	0,1
2	5,0	2	0,2	3	0,3
3	6,2	1	0,1	4	0,4
4	7,3	1	0,1	5	0,5
5	10,5	3	0,3	8	0,8
6	12,0	1	0,1	9	0,9
7	14,0	1	0,1	10	1



Aus der Tabelle liest man $\widehat{F}(13) = 0,9$ ab. D.h. 90% der Pakete wiegen 13 kg oder weniger.

(b) Aus der Urliste errechnet man $\bar{x}^{\text{Med}} = \frac{7{,}3+10{,}5}{2} = 8,9$. Dies lässt sich auch approximativ aus der Grafik ablesen. D.h. die Hälfte der Pakete wiegt höchstens 8,9 kg. Ebenfalls aus der Grafik liest man das 20%-Quantil mit $z_{0{,}2} = 5$ ab. D.h. mindestens 20% der Pakete wiegen 5 kg oder weniger.

(c) Man berechnet:

$$
\begin{aligned}
\bar{x} &= \frac{1}{n}\sum_{i=1}^{n} x_i \\
&= \frac{1}{10}(12 + 10,5 + 7,3 + \ldots + 4,3) \\
&= \frac{85,3}{10} \\
&= 8,53 \\
s_x^2 &= \frac{1}{n}\sum_{i=1}^{n} x_i^2 - \bar{x}^2 \\
&= \frac{1}{10}(12^2 + 10,5^2 + \ldots + 4,3^2) - 8,53^2 \\
&= \frac{830,97}{10} - 72,7609 \\
&= 10,3361
\end{aligned}
$$

$$
\begin{aligned}
s_x &= \sqrt{s^2} \\
&= \sqrt{10,3361} \\
&\approx 3,2150 \\
v_x &= \frac{s}{\bar{x}} \\
&= \frac{\sqrt{10,3361}}{8,53} \\
&\approx \frac{3,2150}{8,53} \\
&= 0,3769
\end{aligned}
$$

- $\bar{x} = 8,53$ bedeutet, dass das Durchschnittsgewicht der Pakete bei 8,53 kg liegt.
- $s_x^2 = 10,3361$ bedeutet, dass die mittlere quadratische Abweichung der Paketgewichte vom Mittelwert (8,53 kg) 10,3361 kg^2 beträgt.
- $s_x = 3,2150$ bedeutet, dass die mittlere standardisierte Abweichung vom Mittelwert 3,2150 kg beträgt.
- $v_x = 0,3769$ bedeutet, dass die mittlere standardisierte Abweichung 37,69% des Mittelwerts beträgt.

(d) Das tatsächliche Gewicht der Pakete liegt 1 kg über dem gemessenen Gewicht. Es handelt sich bei der Berechnung der korrekten Werte um eine lineare Transformation gemäß $y_i := 1 \cdot x_i + 1$. Man erhält somit:

$$\begin{aligned}
\bar{y} &= \bar{x} + 1 \\
&= 8,53 + 1 \\
&= 9,53 \\
s_y^2 &= 1^2 \cdot s_x^2 \\
&= s_x^2 \\
&= 10,3361
\end{aligned}$$

$$
\begin{aligned}
s_y &= s_x \\
&= 3,2150 \\
v_y &= \frac{s_y}{\bar{y}} \\
&= \frac{s_x}{\bar{x}+1} \\
&= \frac{3,2150}{9,53} \\
&= 0,3374
\end{aligned}
$$