Wahrscheinlichkeitsdichte:

$f(x) = P(X=x)$ } Punktwahrscheinlichkeit bei x

Kummulative-W-dichte:

$F(x) = P(X \leq x)$ } Summe der Punktwahrsch. bis x

$F(x) = \int_{-\infty}^{x} f(t)\,dt$

g) Siehe Plot 1-10g.pdf

h) i) $\underbrace{F(1300)} - F(750)$

wir "kennen" aber nur F(1200) und F(1450) wegen der Klassen für Einkommen.

Lösung: Approximiere F(1300) als gewichtetes Mittel von F(1200) und F(1450) unter Annahme einer linearen Approximation

$$\left.\begin{aligned} 1300 - 1200 &= 100 \\ 1450 - 1300 &= 150 \end{aligned}\right\} \text{Gewichte: } \frac{100}{250} = 0.4$$

$$\left.\begin{aligned} F(1200) &= 0.88 \\ F(1450) &= 1 \end{aligned}\right\}$$ Die Funktion geht von 0.88 auf 1 und wir möchten 40% von diesen 0.12 Anstieg

$\Downarrow$

$0.88 + 0.4 \cdot 0.12 = 0.928 \approx F(1300)$

Statistik I Übung (iv)

für $F(750)$ berechnen wir:

$$\left.\begin{array}{l}F(700)=0.48\\ F(900)=0.8\end{array}\right\} 0.48+0.32\cdot\frac{50}{200}=0.56 \approx F(750)$$

somit:

$$P(750\leq Z\leq 1300)=P(Z\leq 1300)-P(Z\leq 750)$$
$$=F(1300)-F(750)$$
$$\approx 0.368$$

Gruppe Woche I

ii) $P(Z>800)=1-P(Z\leq 800)=1-F(800)$

$F(800)\approx 0.48+0.32\cdot\frac{100}{200}=0.64$

$\Rightarrow 1-0.64=0.36$

iii) $F(x)=0.5 \Leftrightarrow x=\underbrace{F^{-1}}_{\text{Quantil, bei 50\% ist es der Median}}(0.5)$

$\left.\begin{array}{l}F(700)=0.48\end{array}\right\}$ nur noch 0.02 "fehlen"

$F(900)=0.8 \Rightarrow$ Differenz 0.32 wenn z um 200 steigt

$\hookrightarrow \frac{200}{32}\cdot 2=12.5$ ist der Anstieg über 2%

$\Rightarrow F^{-1}(0.5)\approx 712.5$

iv) $P(Z>z)\overset{!}{=}0.2 \Leftrightarrow P(Z\leq z)=0.8=F(z)$

$\hookrightarrow$ da $P(Z>z)+P(Z\leq z)=1$ gelten muss

$F(900)=0.8$, somit ist die Antwort 900

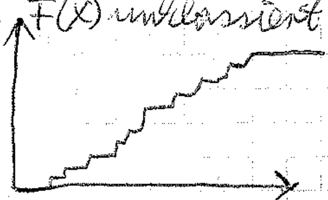
Statistik I Übung - Session 2

Wiederholung:

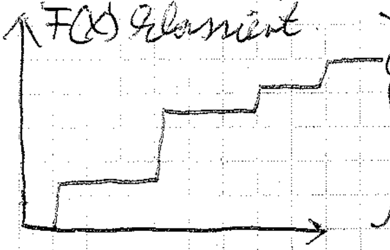
Taschenrechner Casio FX-991ES

Wahrscheinlichkeitsdichte: $f(x) = P(X=x)$

Kummulatiure WKt'funktion: $F(x) = P(X \leq x)$



Skalenniveau:
- Nominal
- Ordinal
- Kardinal < Intervall / Verhältnis / Absolut



} Lineare Approximation für Evaluation der Funktion.

Nr. 1-20

a)

$\#VWL = 5$
$\#Sozi = 5$
$\#BWL = 10$
$\#Pol = 5$

Modus: häufigste Realisierung; Da die Variable nominal skaliert ist gibt es keine Ordnung. Wir sehen jedoch, dass BWL am meisten vorkommt.

Wenn man einen zufälligen Studenten auswählt erwartet man also im Mittel, dass er BWL studiert.

b) Modus: $y_D = 1$

$\#0 = 8$
$\#1 = 11$
$\#2 = 4$
$\#3 = 2$

} 1 kommt am meisten vor

Median: $y_{0.5} = 1 = 50\%$ Quantil

50% Quantil ⇒ 50% der Beobachtungen sind kleiner oder gleich diesem Wert

1 Aufgabe 1-20

1.1 Teil b)

Der Median in diesem Fall wird als die 13. geordnete Beobachtung erhalten. Für eine Stichprobe $\{x_1, \ldots, x_n\}$ mit Umfang n erhält man die Beobachtungsnummer als

$$\frac{n+1}{2} \tag{1}$$

Für $n = 25$ sind genau 12 Beobachtungen größer (und kleiner) als die 13. Beobachtung. In der Aufgabe ist der Median somit $y_D = 1$.

Das Arithmetische Mittel ist definiert als

$$\bar{x} = \frac{1}{n}\sum_{i=1}^{n} x_i \tag{2}$$

Hier beträgt es also:

$$\bar{y} = \frac{8*0 + 11*1 + 4*2 + 2*3}{25} = 1 \tag{3}$$

1.2 Teil c)

$$\bar{y} = \frac{924 + 789 + \ldots + 640}{25} = 808 \tag{4}$$

1.3 Teil d)

Arithmetisches Mittel:

$$\bar{z} = \frac{625*6 + 675*6 + 800*8 + 1050*2 + 1325*3}{25} = 811 \tag{5}$$

Die Aufgabenstellung verlangt nun explizit, dass wir die empirische kummulative Wahrscheinlichkeitsfunktion ($F(z)$) zur Berechnung von Quantilen verwenden. $F(z)$ erhalten wir analog zu Aufgabe 1-10 erhalten, siehe Tabelle 1. Am Punkt 700 beträgt die Funktion genau 0.48. Bis wir 50% erreichen "fehlen" also noch 2%. Innerhalb der Klasse 700-900 steigt die Funktion von 0.48 auf 0.8 an. Die Differenz von 0.32 wird über eine Änderung der Variable Einkommen um 200 erreicht. Der Wert 0.5 wird also approximativ folgenden Einkommenswert erreicht:

$$700 + \frac{200}{32} * 2 = 712.5 = z_{0.5} \tag{6}$$

Klasse	$f(z)$	F(z)
600-650	0.24	0
650-700	0.24	0.24
700-900	0.32	0.48
900-1200	0.08	0.8
1200-1450	0.12	0.88
>1450	0	1

Table 1: Empirische Wahrscheinlichkeitsdichte und kummulative Wahrscheinlichkeitsfunktion

Analog erhalten wir das 75% Quantil in der Klasse 700-900 als:

$$700 + \frac{200}{32} * 27 = 768.75 = z_{0.75} \tag{7}$$

Sowie die 25% und 90% Quantile als:

$$650 + \frac{50}{24} * 1 = 652.08 = z_{0.25} \tag{8}$$

$$1200 + \frac{250}{12} * 2 = 1241.67 = z_{0.9} \tag{9}$$

Nr. 1-22

$\boxed{A}$ a) $X \stackrel{def}{=}$ Auftragshöhe in EUR

↳ Kardinal (Verhältnisskala)

b) ~~Ein~~ Jeder Auftrag ist hier eine Einheit

c) Klassenmittel:

$0 + \frac{20.000 - 0}{2} = 10.000$, etc.

SP-~~Gesamt~~mittel

$$\frac{15 \cdot 10.000 + 30 \cdot 35.000 + 100.000 \cdot 45 + 225.000 \cdot 10}{15 + 30 + 45 + 10}$$

$= 79.500$ [EUR]

$\boxed{B}$ $Y = X / 1.5$

$\bar{y} = 60.000\,\$$

Linearkombination (aka "affine transformation"):

$Z = a + b \cdot X$

Für den Mittelwert:

$\bar{z} = a + b \cdot \bar{x}$

Hier:

$\bar{x} = 1{,}5\,\bar{y} = 90.000$ €

$\boxed{C}$ a) Arithmetisches Mittel $\bar{x}$: Anfällig für Ausreißer

Modus x_D: Robust aber nur gut wenn wir viele Daten oder wenige Ausprägungsmöglichkeiten haben

Median: Sehr robust. Hier wegen des extremen Ausreißers zu empfehlen

$\Rightarrow x_Z = x_{0.5} =$ geordnete Observation Nr. 3

$= 12.000$

Statistik Übung – Session 2 – Teil iii

ii) $\bar{X} = \frac{2000+12000+17000+12000+2000}{5}$
$= 9000$

$\boxed{D}$ $\bar{X}_{Flensburg} = 9000$; 5 Observationen
$\bar{X}_{NY} = 90.000$; 95 Obs
$\bar{X}_{B} = 79.500$; 100 Obs

$\Rightarrow \bar{X}_{gesamt} = \frac{9.000 \cdot 5 + 90.000 \cdot 95 + 79.500 \cdot 100}{5+95+100}$
$= 82.725$

Nr. 1-32
A: 20 sec/Stück B: 30 sec/Stück
C: 60 sec/Stück D: 60 sec/Stück

a) Raten: Verwende harmonisches Mittel
Bsp Geschwindigkeit: Fährt ein Auto 1 Km mit 60 $\frac{Km}{h}$ und einen Km mit 40 $\frac{Km}{h}$ ist die Durchschnittsgeschwindigkeit nicht 50 $\frac{Km}{h}$ sondern 48 $\frac{Km}{h}$

$\bar{X}_H = \frac{n}{\sum_{i=1}^{n} \frac{1}{X_i}}$, n = Stichprobengröße
$X_i, i = \{1, ..., n\}$ = Observationen

$\Rightarrow \bar{X}_H = \frac{4}{\frac{1}{20}+\frac{1}{30}+\frac{1}{60}+\frac{1}{60}} = 34.286 \frac{sec}{Stück}$

b) Bsp Geschwindigkeit: Fährt ein Auto eine Minute mit 60 $\frac{Km}{h}$ und eine mit 40 $\frac{Km}{h}$ ist die durchschn. Geschw. 50 $\frac{Km}{h}$

$\Rightarrow \bar{X}$
A: $\frac{60}{20} = 3 \left[\frac{Stück}{Minute}\right]$ B: $\frac{60}{30} = 2 \left[\frac{Stück}{Minute}\right]$

C & D: $1 \left[\frac{Stück}{Minute}\right]$

Durchschnitt pro Stunde:

$$\frac{3+2+1+1}{4\,[\text{Maschienen}]} \cdot 60\,[\text{Minuten}] = 105\left[\frac{\text{Stück}}{\text{Stunde}}\right]$$

II $Y = \frac{\text{Stück}}{\text{Stunde}}$

a) Durchschnittszeit pro Stück:

$$\frac{20+30+60+60}{4} = 42.5$$

Stück pro Stunde:

$$\frac{60\,[\text{sec}] \cdot 60\,[\text{Min}]}{42.5\left[\frac{\text{sec}}{\text{Stück}}\right]} = 32.5\left[\frac{\text{Stück}}{\text{Stunde}}\right] = \bar{X}$$

b) Durchschnittszeit mit Kontingenten:

$$\frac{1000 \cdot 20 + 500 \cdot 30 + 300 \cdot 60 + 200 \cdot 60}{1000 + 500 + 300 + 200} = 32.5\left[\frac{\text{sec}}{\text{Stück}}\right]$$

$$\frac{60^2}{32.5} = 110.77\left[\frac{\text{Stück}}{\text{Stunde}}\right] = \bar{y}_H$$