1 BeschreibendeStatistik

1.1 Begriffe

]

1.1.1 Beschreibende/Deskriptive Statistik

Beobachtete Daten werden durch geeignete statistische Kennzahlen charakterisiert und durch geeignete Grafiken anschaulich gemacht.

1.1.2 Schließende/Induktive Statistik

Aus beobachtete Daten werden Schlüsse gezogen und diese im Rahmen vorgegebener Modelle der Wahrscheinlichkeitstheorie bewertet.

1.1.3 Grundgesamtheit

Ω: Grundgesamtheit ω:Element oder Objekt der Grundgesamtheit diskret(<30 Ausprägungen), stetig(≥30 Ausprägungen), univariat(p=1), mulivariat(p>1)

1.2 Lagemaße

1.2.1 Modalwerte x_{mod}

Am häufigsten auftretende Ausprägungen (insbesondere bei **qualitativen** Merkmalen)

1.2.2 Mittelwert

R:$mean(x)$
Schwerpunkt der Daten.**Empfindlich**gegemüber Ausreißern.
$\overline{X} = \frac{1}{n}\sum_{i=1}^{n} x_i$

1.3 Median

R:$median(x)$
Liegt in der Mitt der sortierten Daten x_i.
Unempfindlich gegenüber Ausreißern.

$$x_{0.5} = \begin{cases} x_{\frac{n+1}{2}}, \text{falls n ungerade} \\ \frac{1}{2}(x_{\frac{n}{2}} + x_{\frac{n}{2}+1}), \text{falls n gerade} \end{cases} \quad (1)$$

1.4 Streuungsmaße

1.4.1 Spannweite

max x_i - min x_i

1.4.2 Stichprbenverians s^2

R:$var(x)$
Verschiebungssatz:
$s^2 = \frac{1}{n-1}\sum_{i=1}^{n}(x_i - \overline{x}^2) = \frac{1}{n-1}(\sum_{i=1}^{n} x_i^2 - n\overline{x}^2)$ Gemittelte Summe der quadratischen Abweichung vom Mittelwert

1.4.3 Stichprobenstandardabweichung

R:$sd(x)$
$s = \sqrt{s}$ Streuungsmaß mit gleicher Einheit wie beobachteten Daten x_i.$\overline{x}$ minimiert die "quadratische Verlustfunktionöder die Varianz gibt das Mimimum der Fehlerquadrate an.

1.5 p-Quantile

R:$quantile(x,p)$. Teilt die **sortierten** Daten x_i ca. im Verhältnis p: (1-p) d.h. $\hat{F}(x_{(p)}) \approx p$ 1. Quartil = 0.25-Quantil; Median = 0.5-Quantil; 3. Quartil = 0.75-Quartil;

1.6 Interquartilsabstand I

$I = x_{0.75} - x_{0.25}$. Ist ein weiterer Streuungsparameter.

1.7 Chebyshev

$\frac{N(S_k)}{n} > 1 - \frac{1}{k^2}$, für alle k ≥ 1 $\overline{x}$ der Durchschnitt, s > 0 die Stichproben-Standardabweichung von Beobachtungswerten $x_1,...,x_n$. Sei $S_k = \{i, 1 \le i \le n : |x_i - \overline{x}| < k \cdot s\}$; Für eine beliebige Zahl k ≥ 1 liegen mehr als $100 \cdot (1 - \frac{1}{k^2})$ Prozent der Daten im Intervall von $\overline{x} - ks$ bis $\overline{x} + ks$. **Speziell:**Für k = 2 liegen mehr als 75% der Daten im 2s-Bereich um $\overline{x}$. Für k=3 liegen mehr als 89% der Daten im 3s-Bereich um $\overline{x}$. **Komplement Formulierung:** $\overline{S}_k = \{i||x_i - \overline{x}| \ge k \cdot s\}$; $\frac{N(\overline{S}_k)}{n} \le \frac{1}{k^2}$;
Die Ungleichheit lifert nur eine **sehr grobe Abschätzung**, ist aber unabhängig von der Verteilung der Daten. **Empirische Regeln** 68% der Daten im Bereich um $\overline{x} \pm s$. 95% um $\overline{x} \pm 2s$. 99.7% um $\overline{x} \pm 3s$.

1.8 Korrelation

Grafische Zusammenhang zwischen multivariaten Daten y und y durch ein Streudiagramm. Kennzahlen zur Untersuchung des Zusammenhangs:

1.8.1 Empirische Kovarians

R:$cov(x,y)$; $s_{xy} = \frac{1}{n-1}\sum_{i=1}^{n}(x_i - \overline{x})(y_i - \overline{y}) = \frac{1}{n-1}(\sum_{i=1}^{n}(x_i y_i - n\overline{xy})$

1.8.2 Empirische Korrellationskoeffizient r

R:$cor(x,x)$; $r = \frac{s_{xy}}{s_x x_y}$; Näherungsweise lin. Zusammenhang zw. x und y, falls |r| ≈1.

1.8.3 Regressionsgerade y

$y = mx + t$ mit $m = r \cdot \frac{s_y}{s_x}$ und $t = \overline{y} - m \cdot \overline{x}$

2 Wahrscheinlichkeitsrechnung

2.1 Begriffe

Ergebnisraum Ω: Menge aller möglichen Ergebnisse eines Experiments
Elementarereignis $\omega \in \Omega$: einzelnes Element von Ω
Ereignis$E \subseteq \Omega$: beliebige Teilmenge des Ergebnisraums Ω heißt sicheres Ereignis, $\emptyset$ heißt unmögliches Ereignis
Vereinigung $E \cup F$: Ereignis E oder Ereignis F treten ein. $\bigcup_{i=1}^{n} E_i$: mindestens ein Ereignis $E_i$tritt ein.
Schnitt $E \cap F$: Ereignis E und Ereignis F treten ein.
$\bigcap_{i=1}^{n} E_i$ alle Ereignisse E_i treten ein. **Gegenereignis** $\overline{E} = \Omega / E$: Ereignis E tritt nicht ein (Komplement von E)
Disjunkte EreignisseE und F: $E \cap F = \emptyset$

2.2 De Morgan'schen Regeln

$\overline{E_1 \cup E_2} = \overline{E}_1 \cap \overline{E}_2$
$\overline{E_1 \cap E_2} = \overline{E}_1 \cup \overline{E}_2$

2.3 Wahrscheinlichkeit

$0 \le P(E) \le 1$; P(Ω) = 1;
P($\bigcup_{i=1}^{\infty}$) = $\sum_{i=1}^{\infty}$ P(E_i), falls $E_i \cap E_j = \emptyset$ für $i \ne j$

2.3.1 Satz 2.1

P($\overline{E}$) = 1 -P(E)
P($E \cup F$) = P(E)+ P(F) - P($E \cap F$)
(Übungsaufgabe!!! Ergänzen)

2.4 Laplace-Experiment

Zufallsexperimente mit n gleich wahrscheinlichen Elementarereignissen. Dann berechnet sich die Wahrscheinlichkeit P(E) für $E \subseteq \Omega$ aus:
P(E) $= \frac{AnzahlderfrEgnstigenEreignisse}{AnzahldermglichenEreignisse} = \frac{MchtigkeitvonE}{MchtigkeitvonΩ} = \frac{|E|}{\Omega}$ **text**

2.5 Kombinatorik

2.5.1 Allgmeines Zählprinzip

Anzahl der Möglichkeiten für ein k-stufiges Zufallsexperiment mit n_i Varianten im i-ten Schritt: $n_1 \cdot n_2 \cdot ... \cdot n_k$

2.5.2 Permutationen

Anzahl einer n-elementigen Menge n-maliges Ziehen ohne Zurücklgen mit Beachtung der Reihenfolge: **n unterscheidbare Elemente**: $n! = n \cdot (n-1)textbf...2 \cdot 1$
k Klassen mit je n_i **nicht unterscheidbaren Elementen** $n = sum_k^{i=1} n_i$: $\frac{n!}{n_1!\cdot n_2!\cdot ... \cdot n_k!}$

2.5.3 Anzahl k-elementigen Teilmengen einer n-elementigen Menge k-maliges Ziehen aus einer n-elementigen Menge

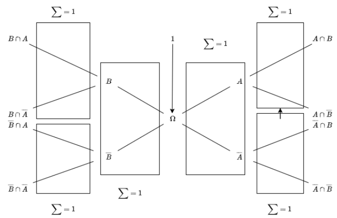
ohne Zurücklegen = $k \le n$.
mit Zurücklegen = $k > n$ möglich.
mit Beachtung der Reihenfolge, ohne Zurücklegen: $\frac{n!}{(n-k)!}$
ohne Beachtung der Reihenfolge, ohne Zurücklegen: $\binom{n}{k} = \frac{n!}{(n-k)!k!}$
mit Beachtung der Reihenfolge, mit Zurücklegen: n^k
ohne Beachtung der Reihenfolge, mit Zurücklegen $\binom{n+k-1}{k}$

2.6 Bedingte Wahrscheinlichkeit

$P(E|F) = P_F(E) = \frac{|E \cap F|}{|F|} = \frac{P(E \cap F))}{P(F)}$

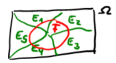
2.6.1 Satz 2.2

$P(E \cap F) = P(E|F) \cdot P(F)$
$P(E \cap F) = P(F|E) \cdot P(E)$



2.6.2 Satz der totalen Wahrscheinlichkeit

Sei $\Omega = \bigcup_{i=1}^{n} E_i$ mit $E_i \cap E_j = \emptyset$ für $i \ne j$ d.h. die Ereignisse bilde eine disjunkte Zerlegung bzw. eine Partition von Ω. Somit gilt:
$P(F) = \sum_{i=1}^{n} P(F \cap E_i) = \sum_{i=1}^{n} P(F|E_i) \cdot P(E_i)$
Summe der Äste des Wahrscheinlichkeitsbaums zu allen Schnitten $F \cap E_i$



2.6.3 Vierfeldertafel

$P(F) = P(F \cap E) + P(F \cap \overline{E})$
$P(\overline{F}) = P(\overline{F} \cap E) + P(\overline{F} \cap E)$

	E	$\overline{E}$	
F	$P(F \cap E)$	$P(F \cap \overline{E})$	$P(F)$
$\overline{F}$	$P(\overline{F} \cap E)$	$P(\overline{F} \cap E)$	$P(\overline{F})$
	$P(E)$	$P(\overline{E})$	1

Satz 2.2 $P(E \cap F)P(E) \cdot P(F|E) = P(F) \cdot P(E|F)$ **Tafel** $= P(F) - P(F \cap \overline{E}) = P(E) - P(\overline{F} \cap E)$; $P(\overline{F}|E) = 1 - P(F|E)$

2.6.4 Formel von Bayes

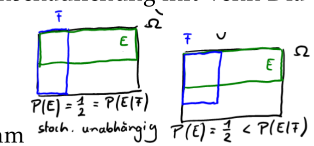
Hilfreich, wenn man man $P(F|E_i)$ kennt, aber nicht $P(E_k|F)$ **Satz 2.4** $P(E_k|F) = \underbrace{\frac{P(F|E_k) \cdot P(E_k)}{\sum_{i=1}^{n} P(F|E_i) \cdot P(E_i)}}$

Nur Nenner!P(F) aus dem Satz der totalen Wahrscheinlichkeit.

2.6.5 Stochastische Unabhängigkeit

Uebung Die Ereignisse E und F heißen (stochastisch) unabhängig, wenn die Information über das Eintreten des einen Ereignisses die Wahrscheinlichkeit für das Eintreten des anderen Ereignisses nicht ändert, d.h. falls
$\underbrace{P(E|F) = P(E) oder P(E \cap F) = P(E) \cdot P(F)}$
$= \frac{P(E \cap F)}{P(F)}$
Es gilt Falls die Ereignisse E, F unabhängig sind, dann sind auch:

$E, \overline{F}$
$\overline{E}, F$
$\overline{E}, \overline{F}$ unabhängig **Bemerkung**

- Stochastische Unabhängigkeit bedeutet nicht notwendigerweise eine kausale Abhängigkeit
- Veranschaulichung mit Venn Diagramm 

- $A, B \ne \emptyset$ und $A \cap B = \emptyset$
  $P(A \cap B) \overset{?}{=} P(A) \cdot P(B)$
  $\emptyset \ne P(A) \cdot P(B)$ da P(A) > 0 und P(B)> 0
  => A, B stochastisch abhängig

3 Zufallsvariable

Abbildung des **abstrakte** Ergebnisraums Ω auf $\mathbb{R}$. Eine Abbildung $X : \Omega \to \mathbb{R}$, $\omega \mapsto X(\omega)$ = heißt Zufallsvariable (ZV). x ∈ R. heißt Realisation der ZV X.

- Diskrete ZV: $X(\Omega) = x_1, ..., x_2 (n \in \mathbb{N})$; z.B. X = "Augensumme beim Würfeln "
- Stetige ZV: $X(\Omega) \subseteq \mathbb{R}$; "z.B. Körpergröße eines Menschen"

3.1 Verteilungsfunktion-allg.

Die Wahrscheinlichkeit P(B) für ein Ereignis B in $\mathbb{R}$ wird zurückgeführt auf die Wahrscheinlichkeit der entsprechenden Ereignisse in Ω. Für jedes $X \in \mathbb{R}$ ist die

Verteilungsfunktion F: $\mathbb{R} \to [0,1]$ einer ZV X definiert durch:
F(x) = P(X ≤ x)

- $0 \leq F(x) \leq 1$
- $\lim_{x \to -\infty} F(X) = 0$ $\lim_{x \to \infty} F(x) = 1$
- monoton wachsend
- $P(X > x) = 1 - F(x)$
- $P(a < X \leq b) = F(b) - F(a)$

3.2 Diskrete ZVs

Für eine diskrete ZV X mit $X(\Omega) = x_1, ..., x_n$ (n endlich oder abzählbar unendlich) ist die Wahrscheinlichkeitsfunktion definiert durch:

$$p(x) = \begin{cases} P(X = x_i), \text{falls } x_i \in X(\Omega) \\ 0, sonst \end{cases} \quad (2)$$

Es gilt:

- $F(x) = (P(X \leq x) = \sum_{x_i \leq x} p(x_i)$
- F(x) ist eine rechtseitig stetige **Treppenfunktion** mit **Sprüngen** bei der Realisation von x_i.

3.3 Stegite ZVs

Stetige ZV X ist die Wahrscheinlichkeitsdichte f $f : \mathbb{R} \to [0, \infty[$ definiert durch
$P(a < X < b) = \int_a^b f(x)dx$
Es gilt:

- $F(x) = P(X \leq x) = \int_{-\infty}^x f(t)dt$ und $F'(x) = f(x)$
- F(x) ist stetig & $P(a < X \leq b) = P(a \leq X \leq b)$ wegen $P(X = a) = 0$

3.4 Verteilungsfunktion

$\int_{\textbf{Untergrenze}}^x$ Es wird normal mit - Integriert.

3.5 Zusammenfassung

3.5.1 Diskrete ZV

- Wahrscheinlichkeitsverteilung p(x) $\sum_i^n p(x_i) = 1 x_i$ ist Realisation der ZV.
- Verteilungsfunktion F(x) ist rechtsseitig stetige **Treppenfunktion. Sprunghöhen:**$P(X = x_i) = F(x_i) - \lim_{x \to x_i -} \neq 0$
- $P(a < X \leq b) = F(b) - F(a) \neq P(a \leq X \leq b)$

3.5.2 Stetige ZV

- Dichtefunktion fx $\int_{-\infty}^{\infty} f(x)dx = 1$
- Verteilungsfunktion $F(x)$ ist stetig mit $F'(x) = f(x); P(X = x_i) = 0$
- $P(a < X \leq b) = F(b) - F(a) = P(a \leq X \leq b) = F(a \leq X < b) = P(a < X < b)$

3.6 Erwartungswert

Der Erwartungswert $E[X] =$ einer ZV X ist der **Schwerpunkt** ihrer Verteilung <u>or</u> der durchschnittliche zu erwartende Wert der ZV.

- diskrete ZV: $E[X] = \sum_{i=1}^n x_i \cdot p(x_i)$
- stetige ZV: $E[X] = \int_{-\infty}^{\infty} x \cdot f(x)dx$

ZV ist konstant. E[X] verhält sich linear.
Eigenschaften von $E[X]$:

- $E[b] = b$
- $E[aX + b] = aE[X] + b$
- $E[\underbrace{X_i + ... + X_n}] = \sum_{i=1}^n E[X_i]$
- $\sum_{i=1}^n x_i$

3.6.1 Satz 3.1

Sei Y = g(X) eine Funktion der ZV X. Dann gilt:

- für diskrete ZV:$E[g(X)] = \sum_{i=1}^n g(x) \cdot p(x_i)$
- für stetige ZV: $E[g(X)] = \int_{-\infty}^{infty} g(x) \cdot f(x)dx$. Das vertauschen von E und g nur bei **linearen** Funktionen möglich. ⇒ g(E[X])

3.7 Varianz

Die Varianz einer ZV X mit µ ist ein quadratisches Streungsmaß. $\sigma^2 = Var[X] = E[\underbrace{(X -)^2}] \overset{\text{falls x stetig}}{=} \int_{-\infty}^{\infty} (x - \mu)^2 \cdot f(x)$
g(X)
Die Standardabweichung $\sigma = \sqrt{Var[X]}$ hat im Gegensatz zur Varianz die gleiche Dimension von die ZV X.

- $Var[b] = 0$
- $Var[aX + b] = a^2 Var[X]$

3.7.1 Satz 3.2

$Var[X] = E[X^2] - (E[X])^2$ **Beim Minuend wird beim Erwartungswert nur das einfach stehende x quadriert <u>nicht</u> f(x)!!!**

3.8 Z-Transformation, Standardisierung

Sei X eine ZV mit µ und σ. Dann ist $Z = \frac{X - \mu}{\sigma} = \frac{x}{\sigma} - \frac{\mu(konstant)}{\sigma}$

3.9 Kovarianz

Eigenschaften:

- $Cov[X, Y] = Cov[Y, X]$
- $Cov[X, X] = Var[X]$
- $Cov[aX, Y] = aCov[X, Y]$

Die Kovarianz zweier ZV (X, Y) ist definiert durch $Cov[X, Y] = E[(X - E[X])(Y - E[Y])$ Die Kovarianz beschreibt die Abhängigkeit zweier ZV X und Y. **Je** stärker diese Korrelieren, **desto** (betragsmäßig) größer ist die Kovarianz. Falls $X, Y \text{stochastisch unabhängig} \Rightarrow Cov[X, Y] = 0$

3.10 Satz 3.3

$Cov[X, Y] = E[XY] - E[X] \cdot E[Y]$

3.10.1 Varianz einer Summe von ZV

- $Var[X_i + ... + X_n] = \sum_{i=1}^n \sum_{j=1}^n Cov[X_i, X_j]$; $Var[X_1 + X_2] = Var[X_1] + Var[X_2] + 2Cov[X_1, X_2]$
- Falls X_i, X_j paarweise unabhängig **!!!**: $Var[X_1 + ... + X_n] = \sum_{i=1}^n Var[X_i]$

3.11 Overview µ σ