1 Deskriptive Statistik

1.1 Begriffe

- Grundgesamtheit Ω
- Element der Grundgesamtheit ω
- diskret (<30) - stetig(≥30)
- univariant (p=1) - mulitvariant (p>1)

1.2 Kenngrößen

- Modalwert: x_{mod}
- Mittelwert: $\bar{x} = \frac{1}{n} \sum_{i=1}^{n}$
- Median: $x_{0.5}$

1.3 Streuungsmaße

- Spannweite: max x_i - min x_i
- Stichprobenvarianz: s^2 = var(x)
  $\frac{1}{n-1} \left(\sum_{i=1}^{n} x_i^2 - n\bar{x}^2\right)$
- Standardabweichung: s = $\sqrt{var(x)}$

1.4 p-Quantile

$$f(x) = \begin{cases} x_{floor(np)+1} & \text{if } n*p \notin \mathbb{N} \\ \frac{1}{2}(x_{np} + x_{np+1}) & \text{if } n*p \in \mathbb{N} \end{cases}$$

1.5 Korrelation

- Empirische Kovarianz: s_{xy}
  $\frac{1}{n-1}\left(\sum_{i=1}^{n}(x_i y_i) - n\bar{x}\bar{y}\right)$
- Empirischer Korrelationskoeffizient: r
  r=$\frac{s_{xy}}{s_x * s_y}$

2 Wahrscheinlichkeitsrechnung

2.1 Begriffe

- Vereinigung E∪F: E oder F
  $\cup_{i=1}^{n} E_i$: >1 Ereignis tritt ein
- Schnitt E∩F: E und F
  $\cap_{i=1}^{n} E_i$: >Alle Ereignisse treten ein
- Gegenereignis $\bar{E}$: nicht E
- Disjunkte Ereignisse E und F: $E \cap F = \emptyset$

2.2 Axiome von Kolmogorov

- $0 \geq P(E) \geq 1$
- $P(\Omega) = 1$
- $P(\cap_{i=1}^{\infty} E_i = \sum_{i=1}^{\infty} P(E_i)$ falls $E_i \cap E_j = \emptyset$

2.3 Bedingte Wahrscheinlichkeit

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

- $P(E \cap F) = P(E|F) * P(F)$
- $P(E \cap F) = P(F|E) * P(E)$

2.4 Formel von Bayes

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

2.5 Stochastische Unabhängigkeit

- $P(E|F) = P(E)$, oder
- $P(E \cap F) = P(E) * P(F)$

3 Zufallsvariablen

3.1 Begriffsklärung

$Eine Abbildung X : \Omega \to \mathbb{N}, \omega \to X(\omega) = x$ Zufallsvariable

- Diskrete ZV: $X(\Omega) = \{x_1, ..., x_n\}$(n∈ ℕ)
- Stetige ZV: X(Ω)⊆ ℕ

3.2 Stetige Zufallsvariablen

Definition: $P(a < X < B) = \int_{a}^{b} f(x)dx$

Es gilt:

- $f(x) \geq 0$
- $\int_{-\infty}^{\infty} f(x)dx = 1$
- $F(x) = P(X \geq x) = \int_{-\infty}^{x} f(t)dt und F'(x) = f(x)$

3.3 Erwartungswert

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

3.4 Varianz und Kovarianz

Varianz: $\sigma^2 = Var[X] = E[(X - \mu)^2]$
Falls stetig: $\int_{-\infty}^{\infty} (x - \mu)^2 * f(x)dx$
Verschiebungssatz: $Var[X] = E[X^2] - (E[X])^2$
Kovarianz: $Cov[X, Y] = E[(X - E[X])(Y - E[Y])]$

3.5 Quantile

Kleinster Wert für den gilt: $F(x_p) \geq p$
Berechnung: $x_p = F^{-1}(p)$

4 Spezielle Verteilungen

4.1 Bernulli Verteilung

- Bei Erfolg 1, bei Misserfolg 0
- Verteilung: $X \sim B_{1,p}$
- Erwartungswert: $E[X] = p$
- Varianz: $Var[X] = p(1 - p)$

4.2 Binomialverteilung

- Wahrscheinlichkeit: $P(X = k) = \binom{n}{k} * p^k * (1 - p)^{n-k}$
- Verteilung: $X \sim B_{n,p}$
- Erwartungswert: $E[X] = np$
- Varianz: $Var[X] = np(1 - p)$

4.3 Hypergeometrische Verteilung

- Wahrscheinlichkeit: $P(X = k) = \frac{\binom{M}{k} * \binom{N}{n-k}}{\binom{M+N}{n}}$
- Verteilung: $X \sim H_{M,N,n}$
- Erwartungswert: $E[X] = n * \frac{M}{M+N}$
- Varianz: $Var[X] = n * \frac{M}{M+N} * (1 - \frac{M}{M+N}) * \frac{M+N-n}{M+N-1}$

4.4 Poisson-Verteilung

- Wahrscheinlichkeit: $P(X = k) = \frac{\lambda^k}{k!} * \exp^{\lambda}$
- Verteilung: $X \sim P_{\lambda}$
- Erwartungswert: $E[X] = \lambda$
- Varianz: $Var[X] = \lambda$

4.5 Gleichverteilung

- Wahrscheinlichkeit: $P(X = x_k) = \frac{1}{n}$
- Verteilung: $X \sim U_{x_1,...,x_n}$
- Erwartungswert: $E[X] = \frac{1}{n} \sum_{k=1}^{n} x_k = \bar{x}$
- Varianz: $Var[X] = \frac{1}{n}(\sum_{k=1}^{n} x_k)^2 - (\bar{x})^2$

4.6 Stetige Gleichverteilung

- Dichte: $f(x) = \frac{1}{b-a}$
- Verteilung: $X \sim U_{a,b}$
- Erwartungswert: $E[X] = \frac{a+b}{2}$
- Varianz: $Var[X] = \frac{(b-a)^2}{12}$

4.7 Normalverteilung

- Dichte: $f(x) = \frac{1}{\sigma\sqrt{2*\pi}} \exp*(-\frac{1}{2}*(\frac{x-\mu}{\sigma})^2))$
- Verteilung: $X \sim N_{\mu,\sigma^2}$
- Erwartungswert: $E[X] = \mu$
- Varianz: $Var[X] = \sigma^2$

4.8 Exponentialverteilung

- Dichte + Verteilung: $f(x) = \lambda * \exp^{-\lambda x}(x \geq 0); F(x) = 1 - (\exp)^{-\lambda x}$
- Verteilung: $X \sim Exp_\lambda$
- Erwartungswert: $E[X] = \frac{1}{\lambda}$
- Varianz: $Var[X] = \frac{1}{\lambda^2}$

4.9 Chiquadrat-Verteilung

- Anwendung: Summen unabhängiger normalverteilter ZV
- Verteilung: $X \sim \chi_n^2$
- Erwartungswert: $E[X] = n$
- Varianz: $Var[X] = 2n$

4.10 t-Verteilung

- Anwendung: Schätz und Testverfahren bei unbekannter Var
- Verteilung: $Y \sim t_n$
- Erwartungswert: $E[Y] = 0$ für n>1
- Varianz: $Var[Y] = \frac{n}{n-2}$ für n>2

5 Zentraler Grenzwertsatz

5.1 General

Seien $X_i (i = 1,...,n)$ ZV, gilt für hinreichend großes n näherungsweise: $\sum_{i=1}^{n} X_i \sim N_{n\mu,n\sigma^2}$.

Wichtig: $\frac{\bar{X}-\mu}{\sigma} * \sqrt{n} \sim N_{0,1}$

Fausteregel für Größe von n:

- $n > 30$, falls unbekannt Verteilung schief ist
- $n > 15$, falls unbekannte Verteilung annähernd symmetrisch
- $n \leq 15$, falls unbekannte Verteilung annähernd normalverteilt

5.2 Stichprobenverteilung

- Stichprobenmittel: $\bar{X} = \frac{1}{n} \sum_{i=1}^{n} X_i$
- Stichprobenvarianz: $S^2 = \frac{1}{n-1} * (\sum_{i=1}^{n} X_i^2 - n\bar{X}^2)$

6 Parameterschätzung

6.1 Konfidenzintervall

- Punktschätzer:
  - Stichprobenmittel: $\bar{X} = \frac{1}{n} \sum_{i=1}^{n} X_i$
  - Stichprobenvarianz: $S^2 = \frac{1}{n.1} \sum_{i=1}^{n} (X_i - \bar{X})^2$
- Intervallschätzer: Konfidenzintervall. das wahren Parameter mit gewisser Wahrscheinlichkeit $(1-\alpha)$ überdeckt
  - Vorgabe einer großen Sicherheit (95

Allgemein: $I =]\bar{X} - \phi^{-1}(1-\frac{\alpha}{2})\frac{\sigma}{\sqrt{n}}, \bar{X} + \phi^{-1}(1-\frac{\alpha}{2})\frac{\sigma}{\sqrt{n}}[$

Aufgabentypen:

- Gesucht n: $\sqrt{n} > 2\phi^{-1}(1-\frac{\alpha}{2})\frac{\sigma}{L}$
- Gesucht 1 - α: $1 - \frac{\alpha}{2} = \phi(\frac{L}{2}\frac{\sqrt{n}}{\sigma})$

7 Hypothesentests

7.1 Hypothesenarten

- Nullhypothese H_0: Angezweifelte Aussage, die widersprochen werden kann (z.B. $H_1 : \mu \neq \mu_0$)
- Gegenhypothese H_1: Gegenteil von H_0 (z.B. $H_1 : \mu \neq \mu_0$

7.2 Signifikanzniveau

- Ablehnungsbereich C: Werte, die für H_1 sprechen und bei Gültigkeit von H_0 mit Wahrscheinlichkeit $\leq \alpha$, dem sog. Signifikanzniveau auftreten. → Fehler 1.Art: H_0 wird verworfen, trotz richtig
- Annahmebereich: Komplement $\bar{C}$ des Ablehnungsbereichs kann nicht abgelehnt werden. → Fehler 2. Art: H_0 wird nicht abgelehnt, obwohl sie falsch ist.

7.3 Klassische Parametertests

Testprobleme:

- Zweiseitiger Test:
  - $H_0 : \mu = \mu_0$ gegen $H_1 : \mu \neq \mu_0$
- Einseitige Tests:
  - $H_0 : \mu \geq \mu_0$ gegen $H_1 : \mu < \mu_0$ bzw.
  - $H_0 : \mu \leq \mu_0$ gegen $H_1 : \mu > \mu_0$

Wird H_0 verworfen, so spricht man von einer signifikanten Schlussfolgerung.

7.4 Gauß-Test

7.4.1 Varianz bekannt

Prüfgröße tg = $\frac{\bar{X}-\mu_0}{\sigma_0} * \sqrt{n}$

H_0	H_1	H_0 ablehnen, falls	p-Wert
$\mu = \mu_0$	$\mu \neq \mu_0$	$\lvert tg \rvert > \phi^{-1}(1-\frac{\alpha}{2})$	$2(1-\phi(\lvert tg \rvert))$
$\mu \leq \mu_0$	$\mu > \mu_0$	$tg > \phi^{-1}(1-\alpha)$	$1-\phi(tg)$
$\mu \geq \mu_0$	$\mu < \mu_0$	$tg < \phi^{-1}(\alpha)$	$\phi(tg)$

7.4.2 Varianz unbekannt

Prüfgröße tg = $\frac{\bar{X}-\mu_0}{S} * \sqrt{n} \quad t_{n-1}$

H_0	H_1	H_0 ablehnen, falls	p-Wert
$\mu = \mu_0$	$\mu \neq \mu_0$	$\lvert tg \rvert > t_{n-1}^{-1}(1-\frac{\alpha}{2})$	$2(1-t_{n-1}(\lvert tg \rvert))$
$\mu \leq \mu_0$	$\mu > \mu_0$	$tg > t_{n-1}^{-1}(1-\alpha)$	$1-t_{n-1}(tg)$
$\mu \geq \mu_0$	$\mu < \mu_0$	$tg < t_{n-1}^{-1}(\alpha)$	$t_{n-1}(tg)$

7.5 p-Wert

Wahrscheinlichkeit, bei Zutreffen von H_0 den beobachteten Wert tg der Prüfgröße oder einen noch stärker von μ_0 abweichenden Wert zu bekommen. Beispiel: p-Wert=0.0114, dann:

- H_0kann für $\alpha = 5\%$ abgelehnt werden,
- für $\alpha = 1\%$ aber nicht

Testentscheidungen anhand des p-Werts:

- p-Wert $< 1\%$: Sehr hohe Signifikanz
- $1\% \leq$p-Wert $< 5\%$: Hohe Signifikanz
- $5\% \leq$p-Wert $\leq 10\%$: Signifikanz
- p-Wert $> 10\%$: Keine Signifikanz

8 Fehlerquellen

Quellarten:

- Diskretierungsfehler
- Modellierungsfehler
- Fehler in Eingangsdaten
- Fehler durch Gleitpunktarithmetik

8.1 Maschinengenauigkeit

ϵ ist die kleinste Zahl x mit $rd(1+x) \neq 1$

Rundungsfehler:

- Absolut: $|rd(x) - x| \leq |x| * \epsilon$
- Relativer: $\frac{|rd(x)-x|}{x} \leq \epsilon$

8.2 Kondition und Stabilität

Numerische Lösung eines Problems:

- x : Exakte Eingangsdaten
- f: Analytische Lösung
- $\hat{x}$: Fehlerbehaftete Eingangsdaten
- $\hat{f}$: Numerisches Lösungsverfahren

Gesamtfehler:
$\rightarrow f(x) - \hat{f}(\hat{x}) = f(x) - f(\hat{x}) + (f(\hat{x}) - \hat{f}(\hat{x})$

8.2.1 Kondition

$cond(x) = |\frac{\text{realiver Fehler im Ergebnis}}{relativerFehlerindenEingabedaten}| = |\frac{\frac{f(\hat{x}-f(x)}{f(x)}}{\frac{\tilde{x}-x}{x}}|$
Schlecht konditioniert, wenn cond >> 1

8.2.2 Fehlerfortpflanzung

- $z = f(x)$
- $\Delta f = f(\tilde{x}) - f(x) \approx \text{f'}(x)\Delta x$
- cond $\approx |\frac{\text{f'}(x)}{f(x)}x|$

9 Interpolation

Im Gegensatz zu Approximation nicht geeignet für verrauschte Daten.

9.1 Polynominterpolation

9.1.1 Klassischer / Vandermonde Ansatz

Ziel : Bestimmung der Koeffizienten $a_0, a_1, ..., a_n$, so dass: $p_n(x_i) = y_i = a_n x_i^n + ... + a_1 x_i + a_0$

In Matrixform: $\begin{pmatrix} x_0^n & ... & x_0^2 & x_0 & 1 \\ x_1^n & ... & x_1^2 & x_1 & 1 \\ x_2^n & ... & x_2^2 & x_2 & 1 \\ ... & ... & ... & ... & ... \\ x_n^n & ... & x_n^2 & x_n & 1 \end{pmatrix} \begin{pmatrix} a_n \\ a_{n-1} \\ ... \\ ... \\ a_0 \end{pmatrix} = \begin{pmatrix} y_0 \\ y_1 \\ ... \\ ... \\ y_n \end{pmatrix}$

Problem : Rechenaufwand für Lösung hoch: $\Theta(n^3)$ und für große n schlecht konditioniert

9.1.2 Ansatz nach Lagrange

$\rightarrow p_n(x) = y_0 L_0(x) + y_1 L_1(x) + ... + y_n L_n(x)$

Beispiel:

i	0	1	2
x_i	-2	3	1
y_i	-15	-5	3

$\rightarrow p_2(x) = y_0 L_0(x) + y_1 L_1(x) + y_2 L_2(x)$

- $L_0(x) = \frac{x-x_1}{x_0-x_1} * \frac{x-x_2}{x_0-x_2} = \frac{x-3}{-2-3} * \frac{x-1}{-2-1} = \frac{1}{15}(x-3)(x-1)$
- $L_1(x) = \frac{x-x_0}{x_1-x_0} * \frac{x-x_2}{x_1-x_2} = \frac{x+2}{3+2} * \frac{x-1}{3-1} = \frac{1}{10}(x+2)(x-1)$
- $L_2(x) = \frac{x-x_0}{x_2-x_0} * \frac{x-x_1}{x_2-x_1} = \frac{x+2}{1+2} * \frac{x-3}{1-3} = -\frac{1}{6}(x+2)(x-3)$

$p_2(x) = -15 * L_0(x) + (-5) * L_1(x) + 3 * L_2(x) = -2x^2 + 4x + 1$
Bemerkungen:

- Vorteil: Keine Neuberechnung, wenn sich nur y-Werte ändern
- Nachteil: Neue Stützpunkte: Funktionen müssen neu berechnet werden
- Rechenaufwand: $\theta((n+1)^2)$

9.1.3 Ansatz nach Newton

$\rightarrow p_n(x) = c_0 + c_1(x-x_0) + ... + c_n(x-x_0)(x-x_1)...(x-x_{n-1})$
$\rightarrow$ Rechenaufwand reduziert sich: $\Theta(n^2)$
Beispiel:

i	0	1	2
x_i	-1	3	1
y_i	-15	-5	3

- x_0, y_0 mit x_1, y_1 : $\frac{-5-(-15)}{3-(-2)} = 2$
- x_1, y_1 mit x_2, y_2 : $\frac{3-(-5)}{1-3} = -4$
- x_0, y_0 bis x_2, y_2 : $\frac{-4-2}{1-(-2)} = -2$

Daraus folgt:

- $c_0 = -15$
- $c_1 = 2$
- $c_2 = -2$

Vorteile:

- Rechenaufwand reduziert sich: $\theta(n^2)$
- Hinzufügen von Stützpunkten ohne großen Aufwand möglich

$p_2(x) = c_0 + c_1(x-(-2)) + c_2(x+2)(x-3) = -15 + 2(x+2) - 2(x+2)(x-3) = -2x^2 + 4x + 1$

9.1.4 Horner-Schema

Klassisch: $p_n(x) = a_n x^n + ... + a_0$
Horner-Schema:
$p_3(x) = a_3 x^3 + a_2 x^2 + a_1 x + a_0 = ((a_3 x + a_2)x + a_1)x + a_0$

9.2 Interpolationsfehler

Falls f hinreichend glatt ist und p_n das eindeutige Interpolationspolynom vom Grad n, dann gilt für den Interpolationsfehler:
$f(x) - p_n(x) = \frac{f^{n+1}(\Theta)}{(n+1)!}(x-x_0)...(x-x_n)$ mit $\Theta \in [x_0; x_n]$

9.3 Chebyshev-Punkte

Stützstellen für besser Interpolation. Erhält man durch orthogonale Projektion von gleichverteilten Punkten auf dem Einheitskreis.
Durch die Verwendung wird der Fehler gleichmäßiger verteilt $\rightarrow$ Konvergenz.

9.4 Spline-Interpolation

= Aus Polynomen zusammengesetzte Funktion. S(x) = $s_0(x) frx_0 \leq x < x_1; S_1(x), frx_1 \leq x < x_2 ...$ Definition Grad k:

- Jede Funktion $S_i(x)$ ist ein Polynom vom Grad ≤ k
- S(x) ist (k-1)-mal stetig differenzierbar

$\underline{Vorteil}$: Nach geschickter Umformung der Gleichung: Rechenaufwand $\Theta(n)$

10 Numerische Integration

10.1 Trapezregel

→ Trapeze zwischen Punkte machen zur Hilfe.
Für Teilintervalle mit der gleichen Länge: h = $\frac{b-a}{n}$
Formel: $T_n = h(\frac{f(x_0)}{2} + f(x_1) + ... + f(x_{n-1}) + \frac{f(x_n)}{2})$

10.2 Simpson-Regel

→ Näherung mit kubischen Parabeln.
Voraussetzung: Gerade Anzahl an Parabeln.
Für 2n Teilintervalle mit gleicher Länge h = $\frac{b-a}{2n}$:

- $S_2 = \frac{h}{3}(f(x_0) + 4f(x_1) + 2f(x_2) + 4f(x_3) + f(x_4))$
- $S_3 = \frac{h}{3}(f(x_0) + 4f(x_1) + 2f(x_2) + 4f(x_3) + 2f(x_4) + 4f(x_5) + f(x_6))$

10.3 Fehler der Quadratur

Ordnung einer Integrationsregel: Eine Integrationsregel hat Ordnung p, wenn sie für Polynome vom Grad ≤ p-1 exakte Werte liefert.
Beispiele:

- Ordnung Trapezregel T_1: 2 (Exakt für Polynome von Grad ≤ 1
- Ordnung Newton-Cotes Regeln: mindestens Ordnung k+1

10.3.1 Grenzen der Newton-Cotes-Regeln

- Bei Verwendung vieler äquidistanter Knoten treten die bekannten Probleme von Interpolationspolynomen höheren Grades auf → Gewichte werden negati, also Verfahren instabil
- Die sog. geschlossenen Newton-Cotes-Regeln machen Funktionsauswertungen an den Grenzen des Intervalls erforderlich → Problem mit Singularitäten
- Die Newton-Cotes-Regeln erreichen aufgrund der äquidistanten Knoten nicht die größtmögliche Ordnung

10.4 Gauß-Quadratur

Idee: Wähle die Knoten t_j und Gewichte α_j so, dass man ein Verfahren möglichst großer Ordnung p erhält.
$\underline{Bedingung:}$
$\int_0^1 p_r(t)dt = \sum_{j=0}^{k} \alpha_j p_r(t_j)$ für alle Polynome vom Grad $\leq p-1$
→ Ordnung p = 2k + 2