4 Aufbereitung von Stichproben

Eine Stichprobe kann sich aus wiederholten Simulationsergebnissen mit unterschiedlicher Seed oder Messwerten zusammensetzen. Konfidenzintervalle geben an, wie genau die Stichprobe den wahren Mittelwert schätzt. Standardabweichungen und Quantile geben an, wie stark eine Stichprobe streut. Eine Reihe von Stichprobenmittelwerten, die von einem Parameter abhängen, kann in Abhängigkeit von diesem aufgetragen werden. Durch geschickte Achsenskalierung kann man evtl. die Art der Abhängigkeit ablesen. Bei Stichproben, die von mehreren Parametern abhängen, können Pivot-Tables bei der Ergebnisaufbereitung helfen.

4.1 Wiederholung empirischer Schätzer (LK 4.4)

Gegeben: n *statistisch unabhängige* und identisch verteilte (independent and identically distributed, IID) Zufallsvariablen $X_1, X_2, \ldots, X_n$ mit Mittelwert E[X] und Varianz VAR[X]

- Schätzer für E[X]: empirischer Mittelwert: $\overline{X}(n) = \frac{\sum_{i=1}^{n} X_i}{n}$
- Schätzer für VAR[X]: empirische Varianz: $S^2(n) = \frac{\sum_{i=1}^{n} \left[X_i - \overline{X}\right]^2}{n-1}$

4.2 Güte der Mittelwertwertschätzung (LK 4.6)

- **Starkes Gesetz der großen Zahlen (aus der Wahrscheinlichkeitstheorie):**

$\overline{X}(n) \xrightarrow[n\to\infty]{} E[X]$ mit Wahrscheinlichkeit 1.

- Problem: $n \to \infty$ nicht simulierbar.
- Wie stark schwankt der Schätzer für endliches n?

$$\mathrm{VAR}\left[\overline{X}(n)\right] = \mathrm{VAR}\left[\frac{1}{n}\sum_{i=1}^{n} X_i\right] = \frac{1}{n^2}\mathrm{VAR}\left[\sum_{i=1}^{n} X_i\right] \overset{\substack{\text{IID}\\ \downarrow}}{=} \frac{1}{n^2}\sum_{i=1}^{n}\mathrm{VAR}[X] = \frac{1}{n}\mathrm{VAR}[X] \tag{4.1}$$

$$\text{Schätzwert für } \mathrm{VAR}\left[\overline{X}(n)\right] = \frac{S^2(n)}{n} = \frac{\sum_{i=1}^{n}\left[X_i - \overline{X}\right]^2}{n\cdot(n-1)} \tag{4.2}$$

4.3 Standard-Normalverteilung (LK 4.5)

- Die Verteilungsfunktion der **Standard-Normal-Verteilung** mit Mittelwert 0 und Standardabweichung 1 ist gegeben durch:

  $$\Phi(z) = \frac{1}{\sqrt{2 \cdot \pi}} \int_{-\infty}^{z} e^{-y^2/2} dy \tag{4.3}$$

  Sie ist analytisch nicht darstellbar und kann nur numerisch gelöst werden.
  $\rightarrow$ Werte aus Tafelwerken

- Transformation einer normal verteilten Zufallsvariable Y mit E[Y] und VAR[Y] in eine standard-normal verteilte Zufallsvariable Z: $Z = \frac{Y - E[Y]}{\sqrt{VAR[Y]}}$

  $$\tag{4.4}$$

- **Zentraler Grenzwertsatz:**
  Seien $X_1, \ldots, X_n$ IID beliebig verteilte Zufallsvariablen mit Mittelwert E[X] und VAR[X]>0. Wir definieren $Y = \overline{X}(n)$ als ihren Mittelwert und führen eine neue Zufallsvariable $Z_n = \frac{\overline{X}(n) - E[X]}{\sqrt{VAR[X]/n}}$, die einer Transformation von Y entspricht. Die Verteilungsfunktion der neuen Zufallsvariablen sei mit $F_n(z)$ bezeichnet. Dann gilt:

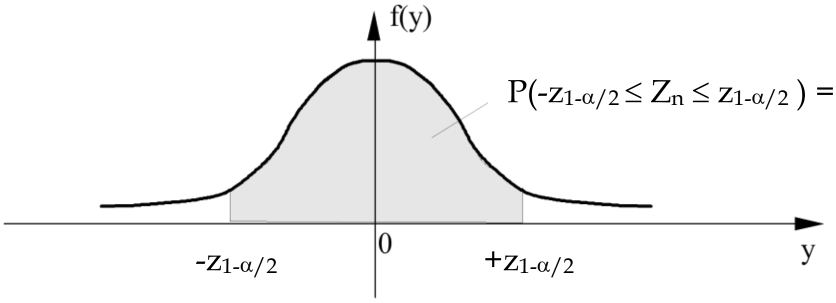
  $$F_n(z) \xrightarrow[n \to \infty]{} \Phi(z). \tag{4.5}$$

  Andere Formulierung: Für hinreichend großes n ist der Mittelwert von n IID Zufallsvariablen normal verteilt mit Mittelwert E[X] und Varianz VAR[X]/n.

4.4 Konfidenzintervalle für Mittelwerte (LK 4.5)

- Es gilt $VAR[X] \approx S^2(n)$ und somit $Z_n \approx \dfrac{\overline{X}(n) - E[X]}{\sqrt{S^2(n)/\,n}}$
- Für hinreichend großes n gilt für den transformierten Mittelwert Z_n
  - $P\left(-z_{1-\alpha/2} \le Z_n \le z_{1-\alpha/2}\right) \approx 1-\alpha$ (4.6)

    wobei $z_{1-\alpha/2}$ das $1-\alpha/2$-Quantil der Standard-Normalverteilung ist.



  - Aussage: Mit einer *statistischen Sicherheit (=Wahrscheinlichkeit)* von $1-\alpha$ liegt eine Stichprobe Z_n im Intervall $[-z_{1-\alpha/2}; +z_{1-\alpha/2}]$.
  - Problemstellung ist aber i. d. R. umgekehrt: Stichproben sind gegeben und Aussage über E[X] wird gesucht.

- Durch Umformung

$$P\left(-z_{1-\alpha/2} \le \frac{\overline{X}(n) - E[X]}{\sqrt{S^2(n)/\,n}} \le z_{1-\alpha/2}\right) \approx 1-\alpha \quad (4.7)$$

$$P\left(\overline{X}(n) - z_{1-\alpha/2} \cdot \sqrt{S^2(n)/n} \le E[X] \le \overline{X}(n) + z_{1-\alpha/2} \cdot \sqrt{S^2(n)/n}\right) \approx 1-\alpha$$

$$l(n,\alpha) = \overline{X}(n) - z_{1-\alpha/2} \cdot \sqrt{S^2(n)/n} \quad (4.8)$$

$$u(n,\alpha) = \overline{X}(n) + z_{1-\alpha/2} \cdot \sqrt{S^2(n)/n}$$

  - [l(n, α); u(n, α)] wird **Konfidenzintervall** oder Vertrauensintervall genannt.
  - $1-\alpha$ ist das **Konfidenzniveau** (confidence level) und α das **Signifikanzniveau** (level of significance)
  - **Aussage:** $1-\alpha$ aller empirisch gewonnenen Konfidenzintervalle enthalten den wahren Wert E[X].

4.5 Konfidenzintervalle für Mittelwerte bei kleinem Stichprobenumfang (LK 4.5)

- Für einen kleinen Stichprobenumfang mit n Elementen nimmt man das 1-α/2-Quantil $t_{n-1,\,1-\alpha/2}$ der **Student-t Verteilungsfunktion** (Gosset, 1908) mit „n-1 Freiheitsgraden" zur Berechnung des Konfidenzintervalls
  - Verteilungsdichtefunktion

$$f_{n-1}(t) = \frac{1}{\sqrt{(n-1)\pi}} \cdot \frac{\Gamma\left(\frac{n}{2}\right)}{\Gamma\left(\frac{n-1}{2}\right)} \cdot \left(1 + \frac{t^2}{n-1}\right)^{-\frac{n}{2}} \tag{4.9}$$

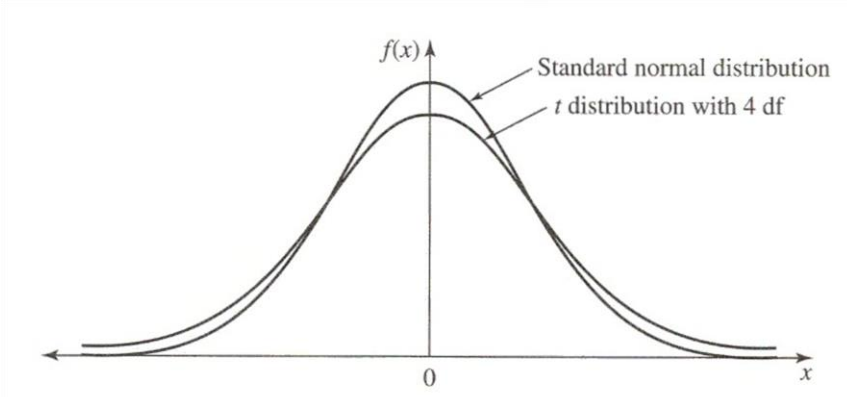


Abbildung 4.1: Die Standard-Normalverteilung hat weniger Varianz als die Student-t-Verteilung mit 4 Freiheitsgraden. (aus Law/Kelton: "Simulation Modeling and Analysis", 3rd Edition, S. 256)

  - Die Student-t-Verteilungen konvergieren gegen die Standard-Normalverteilung für $n \to \infty$.
  - Sie sind flacher und haben eine größere Varianz.
  - Falls die $X_1, \ldots, X_n$ normal verteilt sind, ist das Konfidenzintervall exakt.
  - Je schiefer die Verteilung der Stichprobenwerte ist, desto kleiner ist der Prozentsatz der Konfidenzintervalle, die den wahren Wert E[X] tatsächlich beinhalten (LK Table 4.1, S. 257).
  - Das Konfidenzintervall wird als *Student-t-Vertrauensintervall* bezeichnet.

Tabelle 1: 97,5%-Quantile der Student-t Verteilungen für Konfidenzintervalle mit einem Signifikanzniveau von α=5%.

Stich-proben-umfang n	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**	**20**	∞
$t_{n-1,\,1-\alpha/2}$	12,706	4,303	3,182	2,776	2,571	2,447	2,365	2,306	2,262	2,093	1,960

One Sided	75%	80%	85%	90%	95%	97.5%	99%	99.5%	99.75%	99.9%	99.95%
Two Sided	**50%**	**60%**	**70%**	**80%**	**90%**	**95%**	**98%**	**99%**	**99.5%**	**99.8%**	**99.9%**
df=1	1.000	1.376	1.963	3.078	6.314	12.71	31.82	63.66	127.3	318.3	636.6
2	0.816	1.061	1.386	1.886	2.920	4.303	6.965	9.925	14.09	22.33	31.60
3	0.765	0.978	1.250	1.638	2.353	3.182	4.541	5.841	7.453	10.21	12.92
4	0.741	0.941	1.190	1.533	2.132	2.776	3.747	4.604	5.598	7.173	8.610
5	0.727	0.920	1.156	1.476	2.015	2.571	3.365	4.032	4.773	5.893	6.869
6	0.718	0.906	1.134	1.440	1.943	2.447	3.143	3.707	4.317	5.208	5.959
7	0.711	0.896	1.119	1.415	1.895	2.365	2.998	3.499	4.029	4.785	5.408
8	0.706	0.889	1.108	1.397	1.860	2.306	2.896	3.355	3.833	4.501	5.041
9	0.703	0.883	1.100	1.383	1.833	2.262	2.821	3.250	3.690	4.297	4.781
10	0.700	0.879	1.093	1.372	1.812	2.228	2.764	3.169	3.581	4.144	4.587
11	0.697	0.876	1.088	1.363	1.796	2.201	2.718	3.106	3.497	4.025	4.437
12	0.695	0.873	1.083	1.356	1.782	2.179	2.681	3.055	3.428	3.930	4.318
13	0.694	0.870	1.079	1.350	1.771	2.160	2.650	3.012	3.372	3.852	4.221
14	0.692	0.868	1.076	1.345	1.761	2.145	2.624	2.977	3.326	3.787	4.140
15	0.691	0.866	1.074	1.341	1.753	2.131	2.602	2.947	3.286	3.733	4.073
16	0.690	0.865	1.071	1.337	1.746	2.120	2.583	2.921	3.252	3.686	4.015
17	0.689	0.863	1.069	1.333	1.740	2.110	2.567	2.898	3.222	3.646	3.965
18	0.688	0.862	1.067	1.330	1.734	2.101	2.552	2.878	3.197	3.610	3.922
19	0.688	0.861	1.066	1.328	1.729	2.093	2.539	2.861	3.174	3.579	3.883
20	0.687	0.860	1.064	1.325	1.725	2.086	2.528	2.845	3.153	3.552	3.850
21	0.686	0.859	1.063	1.323	1.721	2.080	2.518	2.831	3.135	3.527	3.819
22	0.686	0.858	1.061	1.321	1.717	2.074	2.508	2.819	3.119	3.505	3.792
23	0.685	0.858	1.060	1.319	1.714	2.069	2.500	2.807	3.104	3.485	3.767
24	0.685	0.857	1.059	1.318	1.711	2.064	2.492	2.797	3.091	3.467	3.745
25	0.684	0.856	1.058	1.316	1.708	2.060	2.485	2.787	3.078	3.450	3.725
26	0.684	0.856	1.058	1.315	1.706	2.056	2.479	2.779	3.067	3.435	3.707
27	0.684	0.855	1.057	1.314	1.703	2.052	2.473	2.771	3.057	3.421	3.690
28	0.683	0.855	1.056	1.313	1.701	2.048	2.467	2.763	3.047	3.408	3.674
29	0.683	0.854	1.055	1.311	1.699	2.045	2.462	2.756	3.038	3.396	3.659
30	0.683	0.854	1.055	1.310	1.697	2.042	2.457	2.750	3.030	3.385	3.646
40	0.681	0.851	1.050	1.303	1.684	2.021	2.423	2.704	2.971	3.307	3.551
50	0.679	0.849	1.047	1.299	1.676	2.009	2.403	2.678	2.937	3.261	3.496
60	0.679	0.848	1.045	1.296	1.671	2.000	2.390	2.660	2.915	3.232	3.460
80	0.678	0.846	1.043	1.292	1.664	1.990	2.374	2.639	2.887	3.195	3.416
100	0.677	0.845	1.042	1.290	1.660	1.984	2.364	2.626	2.871	3.174	3.390
120	0.677	0.845	1.041	1.289	1.658	1.980	2.358	2.617	2.860	3.160	3.373
∞	0.674	0.842	1.036	1.282	1.645	1.960	2.326	2.576	2.807	3.090	3.291

- Beispiel zur Illustration der Genauigkeit von Konfidenzintervallen

Estimated coverages based on 500 experiments

Distribution	Skewness ν	$n = 5$	$n = 10$	$n = 20$	$n = 40$
Normal	0.00	0.910	0.902	0.898	0.900
Exponential	2.00	0.854	0.878	0.870	0.890
Chi square	2.83	0.810	0.830	0.848	0.890
Lognormal	6.18	0.758	0.768	0.842	0.852
Hyperexponential	6.43	0.584	0.586	0.682	0.774

Abbildung 4.2: Konfidenzintervalle wurden 500 Mal für eine endliche Stichprobengröße n berechnet. Die Daten zeigen den Anteil von ihnen, der den wahren Mittelwert tatsächlich enthält. Konfidenzintervalle sind nur für Normal-verteilte Zufallsvariablen exakt. Mit zunehmender Stichprobengröße verlieren sich aber Ungenauigkeiten auf Grund der Schiefe anderer Verteilungen. (aus Law/Kelton: "Simulation Modeling and Analysis", 3rd Edition, S. 257)

4.6 Konfidenzintervalle für Wahrscheinlichkeiten (LK 9.4.1)

- Schätzwert $\hat{p}$ für Wahrscheinlichkeit p $(0 \le p \le 1)$ eines Merkmals A soll simulativ ermittelt werden: n Ereignisse insgesamt, davon k Ereignisse für Merkmal A
- Empirische Wahrscheinlichkeit: $\hat{p} = \frac{k}{n}$ (4.10)
- Man kann Konfidenzintervalle für diese gesuchte Wahrscheinlichkeit berechnen, indem man Konfidenzintervalle zum empirischen Mittelwert von n Bernoulli-Versuchen mit Hilfe der empirischen Varianz $S^2(n)$ nach Gleichung (4.8) berechnet.
- Ziel im Folgenden: Schätzen von Wahrscheinlichkeiten ohne die Berechnung der empirischen Varianz $S^2(n)$
  - Idee
    - Die Anzahl der Ereignisse für das Merkmal A ist binomial verteilt: $k = Binom(n, p)$
    - $Y = \hat{p} = \frac{k}{n} = \frac{Binom(n,p)}{n}$
    - $E[Y] = p$: Der Gesuchte Erwartungswert ist der Parameter p einer Binomialverteilung
    - Da Y auf eine Binomialverteilung zurückgeht, gilt:
      - $VAR[Y] = \frac{1}{n^2} \cdot VAR[Binom(n,p)] = \frac{p \cdot (1-p)}{n}$
      - Y ist annähernd normalverteilt für große n
  - $\hat{p}$ ist der Schätzwert für p
  - $\dfrac{\hat{p}-p}{\sqrt{\dfrac{\hat{p}(1-\hat{p})}{n}}}$ ist approximativ N(0,1)-verteilt (standardnormalverteilt):
  - $P\left(-z_{1-\alpha/2} \le \dfrac{\hat{p}-p}{\sqrt{\dfrac{\hat{p}\cdot(1-\hat{p})}{n}}} \le z_{1-\alpha/2}\right) \approx 1-\alpha$
  - $P\left(\hat{p} - z_{1-\alpha/2} \cdot \sqrt{\dfrac{\hat{p}\cdot(1-\hat{p})}{n}} \le p \le \hat{p} + z_{1-\alpha/2} \cdot \sqrt{\dfrac{\hat{p}\cdot(1-\hat{p})}{n}}\right) \approx 1-\alpha$
  - $\left[\hat{p} - z_{1-\alpha/2}\sqrt{\dfrac{\hat{p}(1-\hat{p})}{n}}\ ;\ \hat{p} + z_{1-\alpha/2}\sqrt{\dfrac{\hat{p}(1-\hat{p})}{n}}\right]$ (4.11)

    ist das approximative Konfidenzintervall mit der statistischen Sicherheit 1-α, wenn $z_{1-\alpha/2}$ das 1-α/2-Quantil der Standardnormalverteilung N(0,1) ist.

4.7 Interpretation von Fehlerbalken

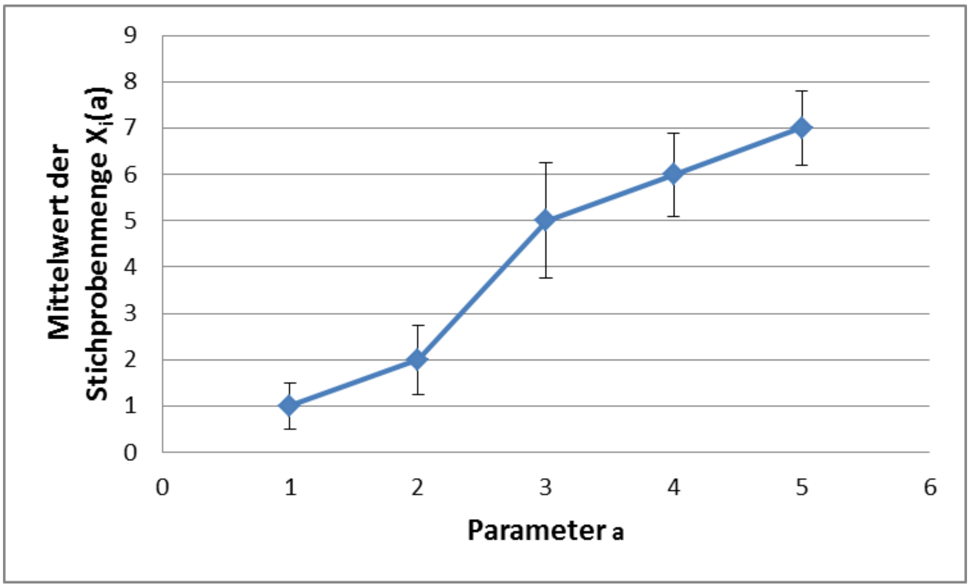


Abbildung 1: In einer Graphik, die Kurven für Mittelwerte E[X(a)] (y-Achse) in Abhängigkeit eines Parameters a (x-Achse) zeigt, werden häufig „Fehlerbalken" (Error Bars) angegeben.

In Graphiken, die beispielsweise Kurven für Mittelwerte $\bar{X}(a)$ in Abhängigkeit eines Parameters a zeigen, werden häufig Fehlerbalken (Error Bars) als Abweichung nach oben und unten angegeben. Diese können unterschiedliche Semantik haben.

- **Minimum und Maximum**
- **Konfidenzintervalle**
  - Aussage über die Genauigkeit des Mittelwertes
  - Keine Aussage über die Schwankung der betrachteten Größe X(a)
  - Mit zunehmender Stichprobengröße kann das Konfidenzintervall verkleinert werden.
- **Standardabweichung**
  - Aussage über die Schwankung der betrachteten Größe X(a)
  - Keine Aussage über die Genauigkeit des Mittelwertes X(a)
  - Standardabweichung hängt von der Verteilung von X(a) ab und kann nicht verkleinert werden.

4.8 Boxplots

Quelle https://de.wikipedia.org/wiki/Boxplot

Boxplots zeigen die Streuung von Stichprobenwerten an.

- Unteres Quartil und oberes Quartil beschreiben eine Box; Höhe der Box ist wird Interquartilabstand (interquartile range (IQR)) genannt
- Median wird in der Box als Querstrich eingezeichnet
- Antennen (Whiskers) verlängern die Box, sie können unterschiedliche Bedeutungen haben
  - Länge der einzelnen Whiskers ist maximal 1.5*IQR, die Whiskers enden aber mit dem letzten in diesem Bereich beobachteten Wert; Werte zwischen 1.5*IQR und 3*IQR werden milde Ausreißer genannt und mit o markiert; Werte außerhalb werden extreme Ausreißer genannt und mit * markiert.
  - Die beiden Whiskers kennzeichnen das 2.5% und 97.5% Qantil
  - Die beiden Whiskers kennzeichnen Minimum und Maximum
- Das arithmetische Mittel (Mittelwert) kann als Stern in den Boxplot aufgenommen werden, ist aber unüblich.

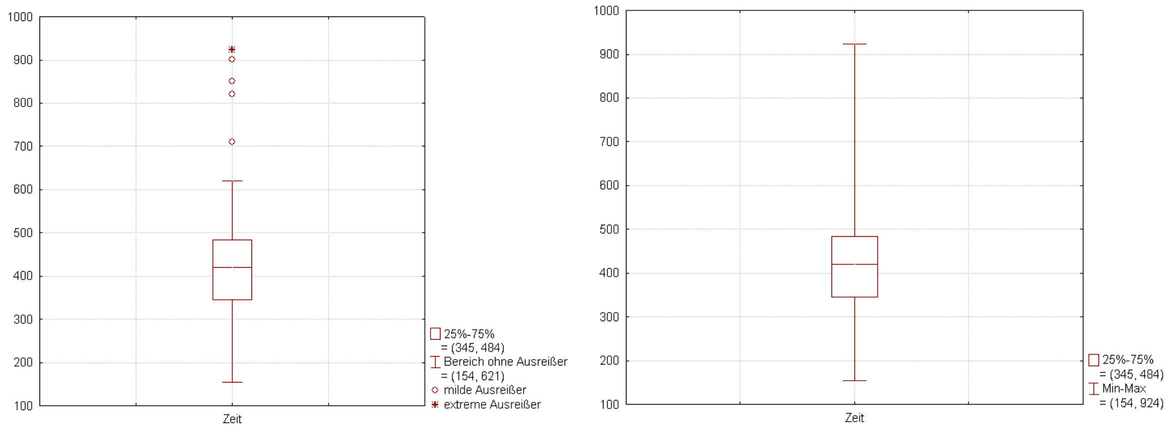


Abbildung 2: Zwei Arten von Boxplots

Quelle: „Box-Plot mit Interquartilsabstand" von Schlurcher - Eigenes Werk. Lizenziert unter CC BY 3.0 über Wikimedia Commons - https://commons.wikimedia.org/wiki/File:Box-Plot_mit_Interquartilsabstand.png#/media/File:Box-Plot_mit_Interquartilsabstand.png; „Box-Plot mit Min-Max Abstand" von Schlurcher - Eigenes Werk. Lizenziert unter CC BY 3.0 über Wikimedia Commons - https://commons.wikimedia.org/wiki/File:Box-Plot_mit_Min-Max_Abstand.png#/media/File:Box-Plot_mit_Min-Max_Abstand.png

4.9 Quantifizieren und Darstellung des Einflusses von Faktoren

Faktoren haben i. A. nichtlinearen Einfluss auf die Systemantwort.

Problem: Wie können logarithmische, polynomielle oder exponentielle Zusammenhänge erkannt und dargestellt werden?

- Idee: Koordinatentransformation durch Skalierung der x- bzw. y-Achse
- Platzierung der Koordinaten
  - Lineare Skalierung: $\hat{x}(x) = x$ bzw. $\hat{y}(y) = y$
  - Logarithmische Skalierung: $\hat{x}(x) = \ln(x)$ bzw. $\hat{y}(y) = \ln(y)$

4.9.1 Skalierung zur Darstellung exponentiellen Wachstums

- Faktor x beeinflusst Systemantwort exponentiell: $y(x) = \exp(\lambda \cdot x)$
- Linear skalierte x-Achse: $\hat{x}(x) = x$
- Logarithmisch skalierte y-Achse: $\hat{y}(x) = \ln(y(x)) = \ln(\exp(\lambda \cdot x)) = \lambda \cdot x = \lambda \cdot \hat{x}(x)$
- Punkte im Koordinatensystem: $(\hat{x}(x), \lambda \cdot \hat{x}(x))$ ergeben eine Gerade mit Steigung λ

4.9.2 Skalierung zur Darstellung logarithmischen Wachstums

- Faktor x beeinflusst Systemantwort logarithmisch: $y(x) = \ln(\lambda \cdot x^{\alpha})$
- Logarithmisch skalierte x-Achse: $\hat{x}(x) = \ln(x)$
- Linear skalierte y-Achse:
  $\hat{y}(x) = y(x) = \ln(\lambda \cdot x^{\alpha}) = \ln(\lambda) + \alpha \cdot \ln(x) = \ln(\lambda) + \alpha \cdot \hat{x}(x)$
- Punkte im Koordinatensystem: $(\hat{x}(x), \ln(\lambda) + \alpha \cdot \hat{x}(x))$ ergeben eine Gerade, die die y-Achse bei $\ln(\lambda)$ schneidet
- Problem: Messpunkte sollen auch bei logarithmischer Skalierung der x-Achse in der Graphik gleichen Abstand voneinander haben, das erfordert einen exponentiellen Abstand der Eingangsparameter im Wertebereich.

- Beispiel
  - Gewünscht: 10 Messwerte im Parameterbereich 1 bis 1000
  - Abstand der Messwerte im logarithmischen Bereich: $\frac{3-0}{10-1}=\frac{3}{9}=\frac{1}{3}$
  - Wähle Messwerte $x_k=10^{\left(\frac{k}{3}\right)}$ mit $k\in[0;9]$

4.9.3 Skalierung zur Darstellung polynomiellen Wachstums (Potenzen)

- Faktor x beeinflusst Systemantwort polynomiell: $y(x)=x^{\lambda}$
- Logarithmisch skalierte x-Achse: $\hat{x}(x)=\ln(x)$
- Logarithmisch skalierte y-Achse: $\hat{y}(x)=\ln(y(x))=\ln\left(x^{\lambda}\right)=\lambda\cdot\ln(x)=\lambda\cdot\hat{x}(x)$
- Punkte im Koordinatensystem: $(\hat{x}(x),\lambda\cdot\hat{x}(x))$ ergeben eine Gerade mit Steigung λ

4.9.4 Skalierung zur Darstellung linearen Wachstums

- Faktor x beeinflusst Systemantwort linear: $y(x)=\lambda\cdot x$
- Linear skalierte x- und y-Achse
  - $\hat{x}(x)=x$, $\hat{y}(x)=y(x)=\lambda\cdot x=\lambda\cdot\hat{x}(x)$
  - Punkte im Koordinatensystem: $(\hat{x}(x),\lambda\cdot\hat{x}(x))$ ergeben eine Gerade mit Steigung λ.
  - Evtl. Abweichungen im Bereich kleiner Parameter sind kaum sichtbar, da bei äquidistantem Abtasten der x-Achse nur große Parameter betrachtet werden.
- Logarithmisch skalierte x- und y-Achse
  - $\hat{x}(x)=\ln(x)$, $\hat{y}(x)=\ln(y(x))=\ln(\lambda\cdot x)=\ln(\lambda)+\ln(x)=\ln(\lambda)+\hat{x}(x)$
  - Punkte im Koordinatensystem: $(\hat{x}(x),\ln(\lambda)+\hat{x}(x))$ ergeben Gerade mit Steigung 1, $e^{\text{Offset zur Winkelhalbierenden auf der y-Achse}}$ gibt Faktor λ an.
  - Evtl. Abweichungen im Bereich kleiner Parameter sind gut sichtbar, da alle Größenordnungen bei äquidistantem Abtasten der x-Achse untersucht werden.

4.10 Organisation von Ergebnissen in PivotTables zur flexiblen Darstellung

Problembeschreibung

- Man hat es oft mit sehr vielen Systemfaktoren zu tun, z.B. mit x1, x2, x3, x4, … und hat für deren Kombinationen Systemantworten (Ergebnisse) ermittelt.
- Die Daten liegen oft in Form einer Ergebnistabelle vor:

x1	x2	x3	R(x1,x2,x3)
1	4	6	24
1	4	7	28
1	5	6	30
1	5	7	35
2	4	6	48
2	4	7	56
2	5	6	60
2	5	7	70
3	4	6	72
3	4	7	84
3	5	6	90
3	5	7	105

- Darstellungsmöglichkeiten für die Systemantwort
  - Zweidimensionales Diagramm: R(x1, x2, x3) auf y-Achse, ein Faktor auf x-Achse, z.B. x1, ein anderer, z.B. x2, wird als Parameter für eine Kurvenschar benutzt.
  - Dreidimensionales Diagramm: R(x1, x2, x3, …) auf y-Achse, zwei Faktoren, z.B. x1 und x2, werden auf den zwei x-Achsen dargestellt.
  - Für die nicht berücksichtigten Faktoren kann dabei
    - entweder ein fester Wert angenommen werden, z.B. x3=6,
    - oder die Systemantwort wird über alle Werte von x3, also x3=6 und x3=7, gemittelt.
  - Problem: oft ist der Haupteinflussfaktor für die Systemantwort a priori noch nicht bekannt und muss noch gefunden werden.
  - Probiere x1, x2, und x3 jeweils als x-Achse für die Darstellung aus!
  - Randbedingung vieler Darstellungsprogramme wie z.B. Excel
    - x- und y-Werte müssen zum Plotten als Vektoren gegeben sein, indem sie in einer Spalte oder Zeile aufeinander folgen.
    - Problem: das ist in der obigen Ergebniszusammenstellung nur für x3 mit festem x1 und x2 gegeben. Beispiel: x3=[6,7] mit R(1,4,[6,7])=[24,28].
  - Für andere Darstellungen muss die Ergebnistabelle umorganisiert werden, das wird durch Pivot-Tables erleichtert.

Benutzung einer Pivot-Tabelle in Excel (Beschreibung für Office 2010)

- Organisation der Daten wie in der Ergebnistabelle
    - Oberste Zeile: Parameterbezeichnungen
    - Restliche Zeilen: Ergebniswerte
- Durchführung des PivotTable-Assistenten (Einfügen → PivotTable)
    - Einlesen der gesamten Tabelle inklusive Überschrift
    - Konfiguration des PivotTables
        - Es erscheint eine PivotTable-Feldliste
        - Ziehe den Faktor für die x-Achse (im Beispiel x1) in das Zeilenbeschriftungen-Feld
        - Ziehe den Faktor für die Kurvenschaar (im Beispiel x2) in das Spaltenbeschriftungen-Feld
        - Ziehe die restlichen Faktoren (im Beispiel x3) in das Berichtsfilter-Feld
        - Ziehe die Systemantwort (im Beispiel R(x1, x2, x3)) in das Werte-Feld
        - Linksklick auf Wertfeldeinstellungen und stelle von Summenwert auf den Mittelwert von R(x1, x2, x3) um!

x3	(Alle)

Mittelwert von R(x1,x2,x3)	x2		
x1	4	5	Gesamtergebnis
1	26	32,5	29,25
2	52	65	58,5
3	78	97,5	87,75
Gesamtergebnis	52	65	58,5

- In den Spaltenbeschriftungen stehen keine Einzelwerte sondern die Durchschnitte über alle sonstigen Faktoren in den Seitenfeldern, man kann aber auch auf einen bestimmten Wert umstellen. (Das geht über die Spaltenbeschriftungsauswahl)
- Parameter im Berichtsfilter, Zeilenbeschriftungen, Spaltenbeschriftungen können beliebig verschoben werden und somit ist es einfach Vektoren im Wertefeld gegen unterschiedliche Vektoren im Zeilenfeld zu plotten.
- Zum Plotten selber: kopiere die Vektoren aus dem PivotTable und erstelle mit Hilfe des Diagramm-Assistenten eine XY-Graphik! Wähle aus, ob Spalten oder Zeilen gegeneinander geplottet werden sollen.