Mathematik Stochastik

MANIT3 – ZUSAMMENFASSUNG

Rémi Georgiou
ZHAW | HS 2015

MANIT3 Analysis 3 – Stochastik	HS 2015	Zürcher Hochschule für Angewandte Wissenschaften zh aw School of Engineering
Dozent	Josef Gohl	

Inhaltsverzeichnis

MANIT3 Analysis 3 – Stochastik	HS 2015	Zürcher Hochschule für Angewandte Wissenschaften zh aw School of Engineering
Dozent	Josef Gohl	

Deskriptive Statistik

Skalierung (Messniveau)

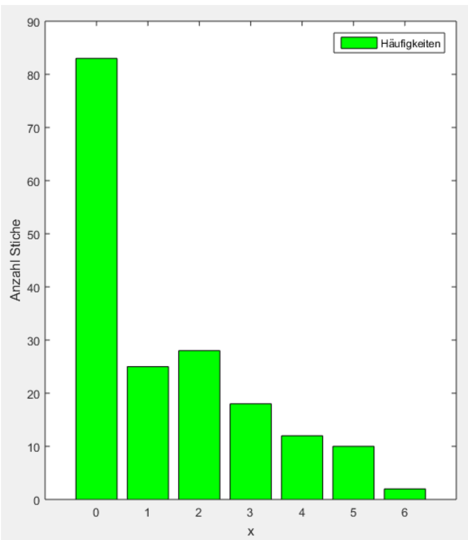
Messniveau	Beschreibung	Beispiele
Nominal	reine Kategorisierung keine Grössenordnung keine Unterschiedsmessung	Wohnort, Studiengang: IT, ET, ST Telefonnummern
Ordinal	Grössenordnung vorhanden, Rangierung möglich Unterschiedsmessung ist nicht messbar	Kleidergrösse: s, m, l Qualifikation: genügend, gut, …
Metrisch (intervall skaliert) **a) Metrisch stetig** **b) Metrisch diskret**	Zahlenskala, insbesondere Grössenunterschiede messbar Grössen sind beliebig genau messbar Es sind nur bestimmte Werte vorgesehen	Gewicht, Einkommen, Anzahl V(t): 574.8 km/h Mathenote, Kinderzahl

Nominal und Ordinal wird zu Kategoriell zusammengefasst.

Graphische Darstellung der Daten:
Kategoriell → Balkendiagramm
Metrisch → Klassenbildung, Histogramm

Kategorielle Daten

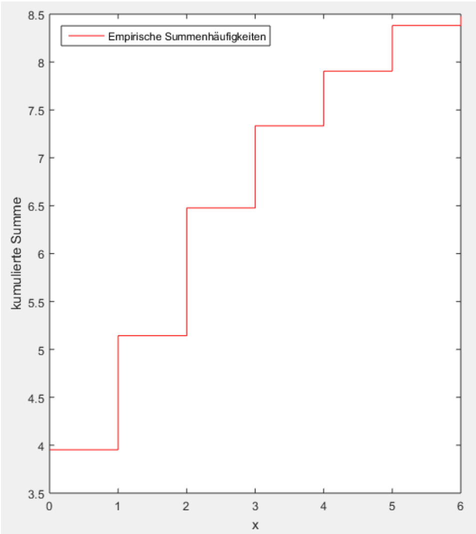
N: Anzahl	f: relative Häufigkeit $f = h/N$
x: Merkmalswert	H: kumulierte Häufigkeit (absolut)
h: Häufigkeit (absolut)	F: kumulierte relative Häufigkeit



Darstellung der Häufigkeitsverteilung mit einer Häufigkeitstabelle und einem **Stabdiagramm**.
Häufigkeiten $\boldsymbol{hi}$ oder $\boldsymbol{fi}$.

x	hi
0	83
1	25
2	28
3	18
4	12
5	10
6	2
Summe	

Die kumulierten Häufigkeiten werden als (empirische) **Summenkurve** $\boldsymbol{Fi}$ dargestellt. → unstetige Sprungfunktion



x	hi
0	83
1	25
2	28
3	18
4	12
5	10
6	2
Summe	

Matlab:
s = cumsum(hi)/sum(x)

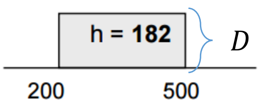
Stetige Merkmale, Dichtebegriff

Einteilung der Daten in **Klassen** (x-Werte von … bis).

X von……bis unter	hi
100 - 200	35
200 - 500	182
500 - 800	317
800 - 1000	84
1000 - 1500	132

Die Häufigkeit h lässt sich **nicht als Stab** über der Klasse darstellen, sondern man stellt die Häufigkeit als Rechteckfläche über der Klassenbreite Δx dar: Die Höhe des Rechtecks ist die (Häufigkeits-) **Dichte D**.
Die Dichte ist eine spezifische Häufigkeit, sie gibt die Häufigkeit pro Einheit von x an.

Absolute Häufigkeit h im Diagramm eingezeichnet.

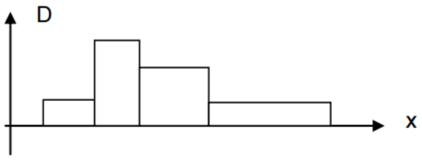


$$Höhe\ des\ Rechtecks = Dichte\ D = \frac{h}{\Delta x} = \frac{182}{500-200} = \frac{182}{300} \approx 0.61$$

0.61 pro Einheit von x in dieser Klasse

Absolute Dichte	Relative Dichte
$D = \frac{h}{\Delta x}$	$d = \frac{f}{\Delta x}$

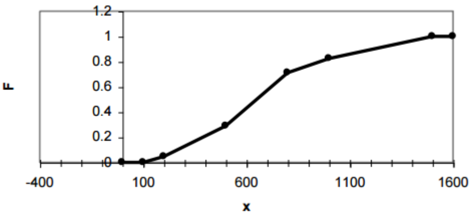
Alle Rechtecke zusammen bilden das Histogramm:



MATLAB:
bar(x-edges,y-values,‘histc‘);

- Das Histogramm ist ein **Flächendiagramm**. Die Flächen sind die Häufigkeiten. Die Rechtecke werden nahtlos aneinander gefügt.
- Bei äquidistanten Klassenbreiten sind die Dichten proportional zu den Häufigkeiten.

Die **Summenkurve (Verteilungsfunktion)** stellt die Summenhäufigkeiten Fi dar. Sie ist eine stückweise lineare Funktion → innerhalb der Klasse i steigt sie um den Betrag fi.

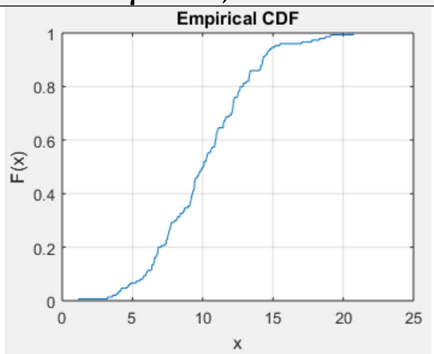
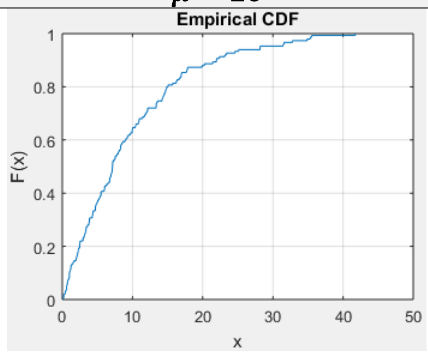
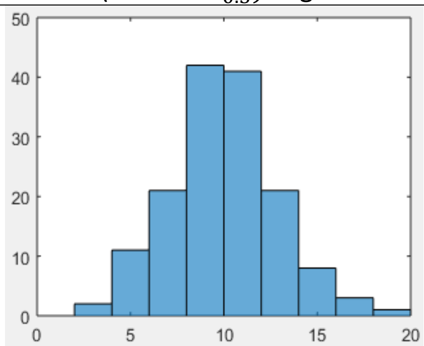
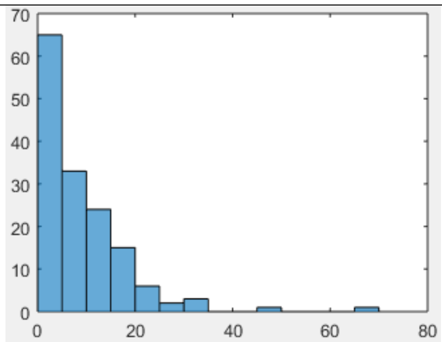



```
MATLAB:
x = 0:length(hi);
s = cumsum(hi)/sum(x);
stairs(x,s);
```

Empirische Verteilungsfunktion $F(x)$

Stellt die **relative** kumulierte Summenhäufigkeit dar. In MATLAB mit cdfplot(x) realisierbar. Besonders wichtig, weil man Teilhäufigkeiten darin ablesen kann und mit wenig Aufwand einen Überblick über die Daten bekommt.

$$F(x) = Fi$$

Symmetrisch → `normrnd(10,3,1,150);` $\mu = 10, \sigma = 3$	**Asymmetrisch** → `exprnd(10,1,150);` $\mu = 10$
 Symmetrische Verteilung Etwa 50 % (Median $x_{0.5}$) liegen unter 10.	 Schiefe Verteilung Median $x_{0.5} < \mu$
 Median ≈ 10 (x-Wert)	 Median ≈ 6 (x-Wert)

Der Median ist der **Mittelwert einer Verteilung** und teilt einen Datensatz in zwei Hälften.

MANIT3 Analysis 3 – Stochastik	HS 2015	Zürcher Hochschule für Angewandte Wissenschaften School of Engineering
Dozent	Josef Gohl	

Quantile

Das α-Quantil mit $\alpha \in [0,1]$ ist der Messwert x, für den $F(x) = \alpha$ gilt. Es wird als x_α bezeichnet.

Quartile:
1. Quartil $= Q1 = x_{0.25}$ 2. Quartil $= Q2 = x_{0.5} =$ **Median** 3. Quartil $= Q3 = x_{0.75}$

Zur Bestimmung des α-Quantils werden die Werte aufsteigend sortiert. Die zu α zugehörige Ordnungsnummer k ist ungefähr $k = \alpha \cdot N$. k ist aber nicht immer ganzzahlig!

Definition α-Quantil:

- $\alpha \cdot N$ ganzzahlig: $k = \alpha \cdot N$ $x_\alpha = \frac{1}{2}(x_k + x_{k+1})$
- $\alpha \cdot N$ nicht ganzzahlig $k = \lceil \alpha \cdot N \rceil$ $x_\alpha = x_k$
  ganzzahlig aufrunden: ceil()

Beispiel:
$X = [4\ 4\ 4\ 4\ 5\ \ 5\ 7\ 7\ 7\ 7\ \ 9\ 13\ 13\ 15]$ $N = 14$
$\alpha = 0.8$: $0.8 \cdot N = 0.8 \cdot 14 \Longrightarrow k = 12$ $x_{0.8} = X_{12} = 13$
$\alpha = 0.5$: $0.5 \cdot N = 0.5 \cdot 14 \Longrightarrow k = 7$ $x_{0.5} = \frac{1}{2} \cdot (7 + 7) = 7$

Klassierte Daten

Median $= X_{0.5}$ durch Interpolation
Bsp: $N = 2000$

x von...bis unter...	hi	Dichte	rel. Summen-häufigkeit Srel (Fi)
300-500	$\Delta x = 200$ 2000*0.12=240	$240/\Delta x = 1.2$	0.12
500 - 600	$\Delta x = 100$ 400-240=160	$160/\Delta x = 1.6$	0.20
600 - 900	$\Delta x = 300$ 1260-400=860	$860/\Delta x = 2.9$	0.63
900 -1300	$\Delta x = 400$ 1620-1260=360	$360/\Delta x = 0.9$	0.81
1300-1800	$\Delta x = 500$ 1840-1620=220	$220/\Delta x = 0.44$	0.92
1800 - 2500	$\Delta x = 700$ 2000-1840=160	$160/\Delta x = 0.23$	1.00

$$\frac{0.63 - 0.2}{900 - 600} = \frac{\mathbf{0.5} - 0.2}{x - 600}$$

$$\Rightarrow x = 809.3023$$

Hi

2000*0.12=240
2000*0.2=400
2000*0.63=1260
2000*0.81=1620
2000*0.92=1840
2000*1.00=2000

$$Dichte\ Di = \frac{hi}{\Delta x}$$

$$Dichte\ di = \frac{fi}{\Delta x}$$

Weiteres Beispiel:

X von……bis unter	hi	fi	Hi	Fi	Δx	Di	di
100 - 200	35	0.047	35	0.047	100	0.35	0.0005
200 - 500	182	0.243	217	0.289	300	0.61	0.0008
500 - 800	317	0.423	534	0.712	300	1.06	0.0014
800 - 1000	84	0.112	618	0.824	200	0.42	0.0006
1000 - 1500	132	0.176	750	1.000	500	0.26	0.0004

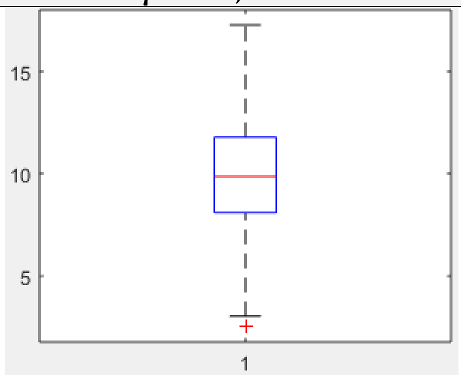
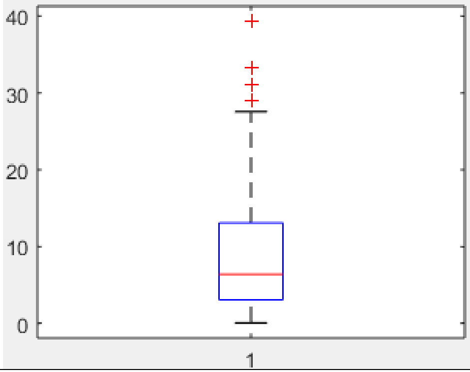
Mittelwert
Den Mittelwert berechnet man mit Hilfe der mit den Häufigkeiten gewichteten Klassenmitten:

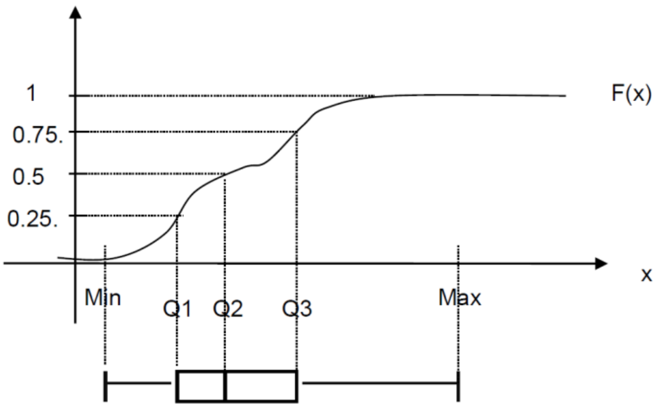
$$\bar{x} = \frac{1}{2000}(450 \cdot 240 + 550 \cdot 160 + 750 \cdot 860 + 1100 \cdot 360 + 1550 \cdot 220 + 2150 \cdot 160) = 961$$

Als **Modus** kann man die Mitte der Klasse mit der grössten Dichte nehmen, hier: $Modus = 750$.

Boxplot-Darstellung

Der Median ist rot dargestellt. Der untere Box-Rand ist das 1. Quartil ($Q1 = x_{0.25}$). Der obere Box-Rand ist das 3. Quartil ($Q3 = x_{0.75}$).

Symmetrisch → `normrnd(10,3,1,150);` $\boldsymbol{\mu = 10, \sigma = 3}$	**Asymmetrisch →** `exprnd(10,1,150);` $\boldsymbol{\mu = 10}$
	



Streumasse

Quadrierte Summen S_{xx}, S_{yy}, S_{xy}:

$$S_{xx} = \sum_{j=1}^{n}(x_j - \bar{x})^2 = \sum_{j=1}^{n}(x_j^2) - n \cdot \bar{x}^2$$

$$S_{xy} = \sum_{j=1}^{n}(x_i - \bar{x})(y_i - \bar{y}) = \sum_{j=1}^{n}(x_j \cdot y_j) - n \cdot \bar{x} \cdot \bar{y}$$

Stichprobenvarianz s^2 (Schätztheorie):

$$s^2 = \frac{1}{n-1} \cdot \sum_{j=1}^{n}(x_j - \bar{x})^2 = \frac{1}{n-1} \cdot \left(\sum_{j=1}^{n}(x_j^2) - n \cdot \bar{x}^2\right)$$

Bei mehreren Variablen:

$$s_x^2 = \frac{S_{xx}}{n-1} = \frac{1}{n-1} \cdot \sum_{j=1}^{n}(x_j - \bar{x})^2 = \frac{1}{n-1} \cdot \left(\sum_{j=1}^{n}(x_j^2) - n \cdot \bar{x}^2\right)$$

Und somit $S_{xx} = s_x^2 \cdot (n-1)$

Standardabweichung s:

$$s = \sqrt{\frac{1}{n-1} \cdot \sum_{j=1}^{n}(x_j - \bar{x})^2} = \sqrt{\frac{1}{n-1} \cdot \left(\sum_{j=1}^{n}(x_j^2) - n \cdot \bar{x}^2\right)} = \sqrt{s^2}$$

Varianz σ^2:

$$\sigma^2 = \frac{1}{n} \cdot \sum_{j=1}^{n}(x_j - \bar{x})^2 = \frac{1}{n} \cdot \left(\sum_{j=1}^{n}(x_j^2) - n \cdot \bar{x}^2\right)$$

Für praktische Berechnungen kann man die Umformung der Stichprobenvarianzgleichung verwenden. Herleitung der Formel für S_{xx}:

$$S_{xx} = \sum_{j=1}^{n}(x_j - \bar{x})^2 = \sum_{j=1}^{n}(x_j^2 - 2 \cdot x_j \cdot \bar{x} + \bar{x}^2)$$

$$= \sum_{j=1}^{n} x_j^2 - \sum_{j=1}^{n}(2 \cdot x_j \cdot \bar{x}) + \sum_{j=1}^{n} \bar{x}^2 = \sum_{j=1}^{n} x_j^2 - 2 \cdot \bar{x} \underbrace{\sum_{j=1}^{n}(x_j)} + \sum_{j=1}^{n}(\bar{x}^2)$$

$n \cdot \bar{x}$: „n mal Mittelwert“

$$= \sum_{j=1}^{n}\left(x_j^2\right) - 2\cdot\bar{x}\cdot n\cdot\bar{x} + \sum_{j=1}^{n}\left(\bar{x}^2\right) = \sum_{j=1}^{n}\left(x_j^2\right) - 2\cdot n\cdot\bar{x}^2\cdot + n\cdot\bar{x}^2 = \sum_{j=1}^{n}\left(x_j^2\right) - n\cdot\bar{x}^2$$

Berechnung der Summe (aus Standardabweichung und Mittelwert)

$$\sum_{j=1}^{n} x_j^2 = s_x^2\cdot(n-1) + n\cdot\bar{x}^2$$

Kovarianz s_{xy}: Ein Zusammenhangsmass zwischen zugeordneten Datenpaaren (x, y)
Sie stellt ein mittleres Abweichungsrechteck vom Punkt $P(\bar{x}, \bar{y})$ dar.

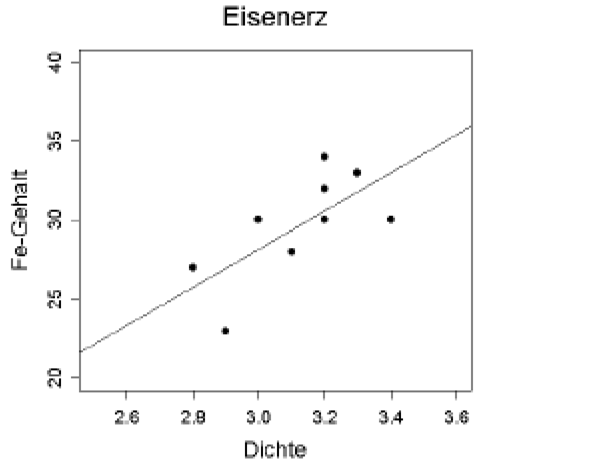
$$s_{xy} = \frac{1}{n-1}\left(\sum_{j=1}^{n}\left(\left(x_j-\bar{x}\right)\cdot\left(y_j-\bar{y}\right)\right)\right) = \frac{1}{n-1}\left(\sum_{j=1}^{n}\left(x_j y_j\right) - n\cdot\bar{x}\cdot\bar{y}\right)$$

MANIT3 Analysis 3 – Stochastik	HS 2015	Zürcher Hochschule für Angewandte Wissenschaften zh aw School of Engineering
Dozent	Josef Gohl	

Lineare Regression

Wir beobachten zwei Variablen x und y. Man interessiert sich dafür, ob zwischen den beiden Variablen eine grössenmässige Beziehung besteht und von welcher Art diese ist. Man fasst die y-Werte als Abhängige von x auf.

z. b. Dichte x und Eisen-Gehalt (%) y



Es ist naheliegend einen linearen Zusammenhang zwischen der Dichte und dem Eisen-Gehalt zu vermuten. Das Bestimmen einer besten Geraden nennt **man lineare Regression**. Diese Gerade heisst **Regressionsgerade** (alter Begriff) oder auch **Ausgleichsgerade** (in der Technik gebraucht).

Regressionsgleichung $y = mx + b$

$\underbrace{y = f(x)}$ + Zufallsfehler

Modell

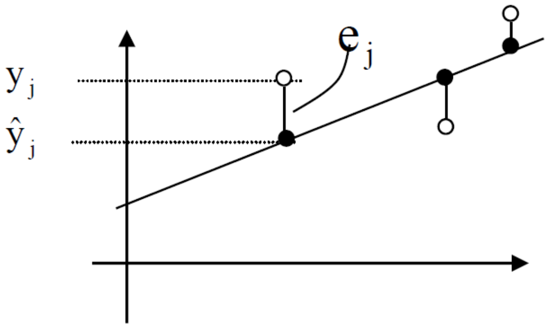
Mit der Geradengleichung $y = mx + b$ berechnet man ideale y-Werte aus den x-Werten.

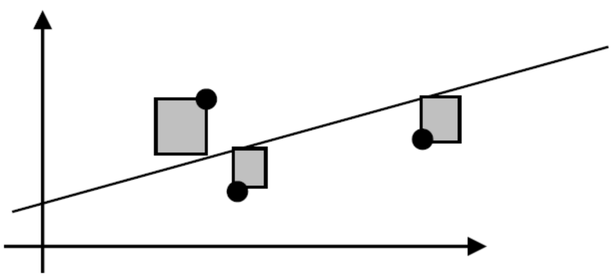
Die Schätzung wird notiert als $\hat{y} = f(x)$

Messwerte oder beobachtete Werte: y_i

Die berechneten Werte werden als erklärte Werte bezeichnet: $\hat{y}_i$

Die Differenzen $y_i - \hat{y}_i = \boldsymbol{e_i}$ sind **Residuen** (Fehler).





Man bestimmt die Gerade so, dass die Summe der **Residuenquadrate** möglichst klein wird:

$$\sum_{i=1}^{n} (e_i^2) = min$$

In Matrixform:

$$\begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix} = \begin{pmatrix} 1 & x_1 \\ \vdots & \vdots \\ 1 & x_n \end{pmatrix} \cdot \begin{pmatrix} \boldsymbol{b} \\ \boldsymbol{m} \end{pmatrix} + \begin{pmatrix} e_1 \\ \vdots \\ e_n \end{pmatrix}$$

$$\bar{y} \quad = \quad \bar{\hat{y}} \quad + \quad \bar{e}$$

Lineares Modell: linear in den Parametern $\boldsymbol{m}$ und $\boldsymbol{b}$.

Gesucht sind Y-Achsenabschnitt b und Steigung m so, dass $\underbrace{\sum_{i=1}^{n}(e_i)}_{\boxed{|e^2|}} = min$

Bestimmen der Geradengleichung

Für die Geradengleichung $y = mx + b$ gilt es, m und b so zu bestimmen, dass die Summe

$$S = \sum_{i=1}^{n} e_i^2 = \sum_{i=1}^{n} (y_i - \hat{y}_i)^2 = \sum_{i=1}^{n} (y_i - mx_i - b)^2$$

minimal wird.

In einem ersten Schritt fassen wir die Terme in der Klammer zusammen und setzen

$z_i = y_i - m \cdot x_i \qquad i = 1, \dots, n$

und schreiben die Summe neu:

$$S = \sum_{i=1}^{n} (z_i - b)^2$$

Diese Summe ist minimal, wenn $b = \bar{z}$ ist. Damit ist $b = \bar{y} - m \cdot \bar{x}$.

$\bar{z}$ ≔ Mittelwert von z
$\bar{x}$ ≔ Mittelwert xon x

Diese Formel deutet bereits die spezielle Lage der Geraden an, denn es folgt daraus:

$$\bar{y} = m \cdot \bar{x} + b$$

Das bedeutet, dass der Punkt $P(\bar{x}, \bar{y})$ (der Schwerpunkt der Punktewolke) die Geradengleichung $y = mx + b$ erfüllt, d.h. P liegt auf der Regressionsgeraden.

Die Gleichung der Regressionsgeraden kann man deshalb auch Punktsteigungsform formulieren:

$$y = m \cdot (x - \bar{x}) + \bar{y}$$

Jetzt muss noch die Steigung m berechnet werden.
Der Ausdruck $y = m \cdot (x - \bar{x}) + \bar{y}$ für b wird in die Summe S eingesetzt.
→ Die rechte Seite ist ein quadratischer Ausdruck in m, den man in die Form $am^2 + bm + c$ bringt.

$$
\begin{aligned}
S = \sum [y_i - m(x_i - \bar{x}) - \bar{y}]^2 &= \sum [(y_i - \bar{y}) - m(x_i - \bar{x})]^2 \\
&= \sum (y_i - \bar{y})^2 + m^2 \sum (x_i - \bar{x})^2 - 2m \sum (y_i - \bar{y})(x_i - \bar{x}) \\
&= S_{xx} \cdot m^2 - 2S_{xy} \cdot m + S_{yy}
\end{aligned}
$$

Die Kurve $S(m)$ ist eine Parabel, S wird am kleinsten im Scheitelpunkt, wir schreiben daher die Scheitelpunktform:

$$S = S_{xx} \cdot \left[m - \frac{S_{xy}}{S_{xx}}\right]^2 + S_{yy} - \frac{S_{xy}^2}{S_{xx}}$$

S wird minimal, wenn $m = \frac{S_{xy}}{S_{xx}}$ gesetzt wird.

Für die Gleichung der Regressionsgeraden $y = mx + b$ gilt:

$$Steigung\ m = \frac{s_{xy}}{s_x^2} = \frac{S_{xy}}{S_{xx}} = \frac{\sum_{i=1}^{n}(x_i - \bar{x})(y_i - \bar{y})}{\sum_{i=1}^{n}(x_i - \bar{x})^2} = \frac{\sum_{i=1}^{n}(x_i \cdot y_i) - n\overline{xy}}{\sum_{i=1}^{n}\left(x_i^2\right) - n\bar{x}^2}$$

$$Y - Achsenabschnitt\ \ b = \bar{y} - m \cdot \bar{x}$$

Die Summe der Residuenquadrate hat den Wert $S = S_{min} = S_{yy} - \frac{S_{xy}^2}{S_{xx}}$

Regressionsgerade $y = mx + b$:
$y = m \cdot (x - \bar{x}) + \bar{y}$

Die **Kovarianz** $\boldsymbol{s_{xy}}$ kann man über die Steigung bestimmen: $\boldsymbol{s_{xy} = m \cdot s_x^2}$

Mittels ***plot(x,y)*** kann in MATLAB eine Scatterplot gezeichnet werden. Die Regressiongerade wird mit dem Befehl ***lsline*** eingezeichnet.

Für die Regression ist eine Matrix X vorzubereiten mit Vektoren als Kolonnen.
n = length(x);
X = [ones(n,1) x‘];
B = regress(y‘, X);

Der Vektor B enthält den Y-Achsenabschnitt b und die Steigung m.

Varianz-Zerlegung und Bestimmtheitsmass

Die erklärten y-Werte sind definiert durch $\boldsymbol{\hat{y}_i = m x_i + b}$ und die Residuen durch $\boldsymbol{e_i = y_i - \hat{y}_i}$.

Beispiel:

x_i	**3**	**5**	**6**	**10**	$\bar{x} = 6$
y_i	**5**	**4**	**8**	**11**	$\bar{y} = 7$
$\hat{y}_i$	4.11	6.04	7	10.85	$\bar{\hat{y}} = 7$
e_i	0.89	-2.04	1	0.15	$\bar{e} = 0$

Offensichtlich gilt: $\bar{\hat{y}} = \bar{y}$ und $\bar{e} = 0$

Diese Eigenschaft ergibt sich aus der kleinste Quadrate Methode (kQM).

$$S_{min} = \sum e_i^2 = \sum (e_i - 0)^2 = S_{ee}$$

Die Residuenvarianz ist deshalb:

$$s_e^2 = \frac{S_{ee}}{n-1}$$

$$S_{yy} = \sum_{i=1}^{n} y_i^2 = 5^2 + 4^2 + 8^2 + 11^2 - 4 \cdot 7^2 = 30$$

$$S_{\hat{y}\hat{y}} = \sum_{i=1}^{n} \hat{y}_i^2 = 4.11^2 + 6.04^2 + 7^2 + 10.85^2 - 4 \cdot 7^2 = 24.096$$

$$S_{ee} = \sum_{i=1}^{n} e_i^2 = 0.89^2 + 2.04^2 + 1^2 + 0.15^2 - 4 \cdot 0^2 = 5.976$$

Man erkennt die Zerlegung $30 = 24.096 + 5.975$

Allgemein gilt die sogenannte **Quadratesummen-Zerlegung**:

$$S_{yy} = S_{\hat{y}\hat{y}} + S_{ee}$$

Teilt man diese Summe durch $n - 1$, so erhält man die **Varianz-Zerlegung**:

$$s_y^2 = s_{\hat{y}}^2 + s_e^2$$

totale Varianz = erklärte Varianz + Fehler-Varianz

Aus diesem Zusammenhang wird eine Vergleichszahl eingeführt, das **Bestimmtheitsmass R^2**.

$$R^2 = \frac{s_{\hat{y}}^2}{s_y^2} \qquad\qquad 0 \le R^2 \le 1$$

Das Bestimmtheitsmass ist der Anteil der erklärten Varianz an der gesamten Varianz. $R^2 = 0.73$ bedeutet, dass 73 % der gesamten Varianz durch die Regression erklärt ist, der Rest ist Zufallsstreuung.

$s_{\hat{y}}^2 = R^2 \cdot s_y^2$
$s_e^2 = (1 - R^2) \cdot s_y^2$

Mit Hilfe der Varianzzerlegung und der Minimumformel der Residuenquadrate erhält man zwei Ausdrücke für R^2:

$$R^2 = \frac{s_{\hat{y}}^2}{s_y^2} = \frac{s_y^2 - s_e^2}{s_y^2} = \frac{s_y^2 - \left[s_y^2 - \frac{s_{xy}^2}{s_x^2}\right]}{s_y^2} = \frac{s_{xy}^2}{s_x^2 \cdot s_y^2}$$

$$\boxed{R^2 = \frac{s_{\hat{y}}^2}{s_y^2} = \frac{s_{xy}^2}{s_x^2 \cdot s_y^2}}$$

Bedeutung von R^2 bezüglich der Regressionsgeraden.

$R^2 = 1$: Dann ist $S_{ee} = 0$, die Punkte liegen exakt auf der Geraden.

$R^2 = 0$: $S_{ee} = S_{yy}$, dies ist dann der Fall, wenn die Regressionsgerade **die Steigung $m = 0$** hat. Dann ist auch $S_{xy} = 0$, d.h. die x- und y-Werte korrelieren nicht.
➔ y lässt sich nicht als Funktion von x modellieren.

Normierung der Kovarianz s_{xy}:

$$r = \mp\sqrt{R^2} = \sqrt{\frac{s_{xy}^2}{s_x^2 \cdot s_y^2}}$$

r heisst **Korrelation** zwischen x und y (genauer: Produktmomenten-Korrelation nach Pearson)

Die **Korrelation r** hat dasselbe Vorzeichen wie die Kovarianz s_{xy} und die Steigung m der Regressionsgeraden und es gilt $r \in [-1,1]$ respektive $R^2 \in [0,1]$.

Die **Kovarianz s_{xy}** kann man über die Varianzen und die **Korrelation r** bestimmen: $\boldsymbol{s_{xy} = r \cdot s_x \cdot s_y}$

Wenn die Steigung der Regressionsgeraden $= 0$ ist, $\Longrightarrow r = s_{xy}$, dann erklärt x nicht y!

Bei standardisierten Daten gilt: $s_x = s_y = 1$

Berechnung der Residuen:
$\hat{y} = x \cdot b$
$e = \bar{y} - \bar{\hat{y}} = y - x \cdot b$

Kombinatorische Formeln

Für die Wahrscheinlichkeitsrechnung kann man die Kombinatorik eingrenzen auf die Fragen nach der Anzahl **Anordnungen** oder der Anzahl **Auswahlen**. Dabei kann geht man von einem Grundprinzip aus, der sog. **Produktregel**.

Produktregel

Das Produkt der Möglichkeiten pro Stufe.

Beispiel 1: Bilden von vierstelligen Passwörtern mit verschiedenen Ziffern; erste Ziffer $\neq 0$.

$$z = 9 \cdot 9 \cdot 8 \cdot 7 = 4536$$

Beispiel 2: Kontrolle von 2 Glühbirnen aus einer 12er-Schachtel Glühbirnen.

$$z = \frac{12 \cdot 11}{2} = 66$$

Anordnungen

Permutation (Vertauschung) – Anordnungsproblem ohne Wiederholungen

Auf wie viele Arten kann man n verschiedene Elemente anordnen?
Gesucht ist die Anzahl aller Anordnungen → Produktregel $\quad z = 4 \cdot 3 \cdot 2 \cdot 1 = 4! = 24$

> n verschiedene Elemente kann man auf $n!$ Arten anordnen.

Permutation (Vertauschung) – Anordnungsproblem mit Wiederholungen

Beispiel: $\quad n = 11$ Elemente $\quad A, A, A, B, B, C, D, E, E, E, E$

$$z = \frac{11!}{3! \cdot 2! \cdot 4!} = 138'600$$

Anwendung bei Binomial-Koeffizienten

Koeffizienten direkt berechnen $\quad (a+b)^n$
Diese 2^n Produkte kommen so häufig vor, wie man die Buchstabenfolgen bestehend aus a und b anordnen kann.

Beispiel: $\quad (a+b)^7 = (a+b) \cdot (a+b) \cdot \ldots \cdot (a+b)$

Für das Produkt $aaaabbb$ erhält man:
$z = \frac{7!}{4! \cdot 3!} = \binom{7}{4} = \binom{7}{3} = 35$ $\quad$ 35 ist der Binomialkoeffizient bei $a^4 b^3$.

Definition der Binomialkoeffizienten: „n tief k“:

$$\binom{n}{k} = \binom{n}{n-k} = \frac{n!}{k! \cdot (n-k)!}$$

$0 \leq k \leq n$

$$(a+b)^5 = \binom{5}{0} a^0 b^5 + \binom{5}{1} a^1 b^4 + \binom{5}{2} a^2 b^3 + \binom{5}{3} a^3 b^2 + \binom{5}{4} a^4 b^1 + \binom{5}{5} a^5 b^0$$

Binomische Formel:

$$(a+b)^n = \sum_{k=0}^{n} \binom{n}{k} a^k b^{n-k}$$

Beispiel: $(2x-5)^7$ wird ausmultipliziert. Welcher Faktor steht bei x^4?

$$\binom{7}{4} \cdot (2x)^4 \cdot (-5)^3 = 35 \cdot 2^4 \cdot (-5)^3 \cdot x^4 = 70'0000 \cdot x^4$$

Zeilensumme im Pascal'schen Dreieck:

$$\sum_{k=0}^{n} \binom{n}{k} = \binom{n}{0} + \binom{n}{1} + \cdots + \binom{n}{n} = 2^n$$

Merke:

$\binom{n}{0} = \binom{n}{n} = 1$ $\qquad$ $\binom{n}{1} = \binom{n}{n-1} = n$ $\qquad$ $\binom{n}{2} = \frac{n \cdot (n-1)}{2}$

Allgemein:

$$\binom{n}{k} + \binom{n}{k+1} = \binom{n+1}{k+1}$$

Auswahlen

Aus einer Stichprobe will man einen Eindruck über die Gesamtheit gewinnen.

Ziehen von $\boldsymbol{k}$ Eementen aus $\boldsymbol{n}$ verschiedenen Elementen ohne Berücksichtigung der Reihenfolge:

$\binom{n}{k}$ verschiedene Auswahlen

Übersicht

Es sind vier Ziehungsarten möglich.

Ohne zurücklegen: $k \leq n$		**Mit zurücklegen: k beliebig**	
Variation k-ter Ordnung **ohne** Wiederholung	Kombination k-ter Ordnung **ohne** Wiederholung	Variation k-ter Ordnung **mit** Wiederholung	Kombination k-ter Ordnung **mit** Wiederholung
Reihenfolge wird berücksichtigt	Reihenfolge spielt keine Rolle	Reihenfolge wird berücksichtigt	Reihenfolge spielt keine Rolle
$\frac{n!}{(n-k)!}$ mit $k \leq n$	$\binom{n}{k}$ mit $k \leq n$	n^k $k > n$ möglich	$\binom{n-1+k}{k}$ $k > n$ möglich

Beispiel 1: Anzahl Siegerlisten der ersten 3 Plätze bei einem Pferderennen mit 7 Pferden.

$$z = \frac{n!}{(n-k)!} = \frac{7!}{(7-3)!} = \frac{7!}{4!} = 7 \cdot 6 \cdot 5 = 210$$

Beispiel 2: Anzahl achtstellige Zahlen bestehend aus lauter ungeraden Ziffern (1,3,5,7,9).

$n = 5, \quad k = 8 \qquad z = 5^8 = 390'625$

Wahrscheinlichkeiten

Klassische Definition nach Laplace

Wurf eines Würfels. Jede Zahl hat dieselbe Wahrscheinlichkeit $\frac{1}{6}$.

$$P(A) = \frac{Anzahl\ für\ das\ Ereignis\ A\ günstige\ Fälle}{Anzahl\ mögliche\ Fälle}$$

Kurz: $p = \frac{g}{m}$

Es gilt immer: $0 \leq P(A) \leq 1$

Beispiel 1: Münze sechsmal werfen. OOOOOO

a) Der dritte Wurf zeigt Zahl

$m = 2, \qquad g = 1 \qquad\qquad p = \frac{g}{m} = \frac{1}{2}$

b) Genau dreimal erscheint Zahl

$m = 2^6 = 64$

$g: zkkzzk$ „3 aus 6 Plätzen" $\frac{6!}{3! \cdot 3!} = \binom{6}{3} = 20 \qquad p = \frac{20}{64} = \frac{5}{16}$

c) Mindestens viermal erscheint Zahl

4x Zahl + 5x Zahl + 6x Zahl $\binom{6}{4} + \binom{6}{5} + \binom{6}{6} = 15 + 6 + 1 = 22 \qquad p = \frac{22}{64} = \frac{11}{32}$

„Entweder, oder"

Beispiel 2: Wurfbilder beim Werfen von zwei Würfeln.

a) Die Augensumme ist 9

$m = 6^2 = 36,\quad g = 4$ Diagonale: (6|3), (5|4), (4|5), (3|6) $p = \frac{4}{36} = \frac{1}{9}$

b) Mindestens eine der Augenzahlen ist eine Primzahl (2,3,5)

Gegenteil von mindestens EINE ist KEINE: Gegenereignis $\bar{g} = 3 \cdot 3 = 9$

$$p = \frac{g}{m} = \frac{m - \bar{g}}{m} = 1 - \frac{\bar{g}}{m} = 1 - \frac{9}{36} = 1 - \frac{1}{4} = \frac{3}{4}$$

Gegenwahrscheinlichkeit

c) Die Differenz der Augenzahlen ist höchstens 2

Diagonalen $g = 6 + 2 \cdot 5 + 2 \cdot 4 = 6 + 10 + 8 = 24$ $p = \frac{24}{36} = \frac{2}{3}$

Beispiel 3: Schachtel mit $N = 12$ Glühbirnen. 4 defekt, 8 intakt, Stichprobe $n = 5$, Ziehen ohne zurücklegen

a) Die Glühbirnen in der Stichprobe sind alle intakt.

$m = \binom{12}{5} = 792,\quad g = \binom{8}{5} = 56$ $p = \frac{56}{792} = 0.071$

b) Höchstens zwei defekte Glühbirnen in der Stichprobe

$g = \binom{8}{5} \cdot \binom{4}{0} + \binom{8}{4} \cdot \binom{4}{1} + \binom{8}{3} \cdot \binom{4}{2} = 56 + 280 + 336 = 672$ $p = \frac{g}{m} = \frac{672}{792} = 0.848$

c) Mindestens ein defekt in der Stichprobe

$$p(min. 1\, defekt) = 1 - p(keine\, defekt) = 1 - \frac{\binom{8}{5}}{\binom{12}{5}} = 0.93$$

Gegenwahrscheinlichkeit

Beispiel Zahlenlotto: Jede Zahl hat die gleiche Chance gezogen zu werden. „Ziehe 6 aus 45".

Anzahl mögliche Tipps: $\binom{45}{6} = 8'145'060$

Wahrscheinlichkeit für 0 Richtige: $\frac{\binom{6}{0} \cdot \binom{45-6}{6}}{\binom{45}{6}} = \frac{3'262'623}{8145'060} = 0.4005$

Allgemeine Formel zur Berechnung von k Richtigen im Lotto:

$$a_k = \binom{6}{k} \cdot \binom{45-6}{6-k}$$

Ein Ereignis ist eine Ergebnis-Konstellation. Ereignis $A \subset \Omega$

$P(A)$ Probabilität des Ereignis A. Die Menge der Ereignisse heisst **Ereignisraum** oder **Wahrscheinlichkeitsraum** $\Omega = \{\omega_1, \omega_2, \omega_3, \omega_4, \dots, \omega_n\}$.

Das sichere Ereignis $A = \Omega$

Das unmögliche Ereignis $A = \{\}$

Axiome der Wahrscheinlichkeit

Jedem Ereignis $A \subset \Omega$ wird eine Zahl $P(A)$ zugeordnet mit den Eigenschaften:

A1:	$0 \leq P(A) \leq 1$
A2	$P(\Omega) = 1$
A3	falls $A \cap B = \{\}$, dann $P(A \cup B) = P(A) + P(B)$

(gemäss A. Kolmogoroff)

Wahrscheinlichkeitsfunktion $A \longrightarrow P(A)$

Diese drei Axiome genügen, wenn der Ereignisraum Ω endlich ist. Für den Falls, dass es unendlich viele Teilmengen gibt, ist noch ein weiteres Axiom notwendig (→ sogenannte Ereignisalgebra).

Zu A3: Prinzip von Inklusion und Exklusion

$P(A \cup B) = P(A) + P(B) - P(A \cap B)$
Additionssatz: $|A \cup B| = |A| + |B| - |A \cap B|$

Folgerungen aus den Axiomen:

1. $\Omega = A \cup \bar{A}$ und $A \cap \bar{A} = \{\}$ $\Longrightarrow P(A \cup \bar{A}) = P(A) + P(\bar{A}) = 1$ $\Longrightarrow P(\bar{A}) = 1 - P(A)$

$P(\bar{A})$ ist die **Gegenwahrscheinlichkeit** von $P(A)$.
Speziell : $P(\{\}) = 1 - P(\Omega) = 0$

2. $A \subseteq B \Longrightarrow P(A) \leq P(B)$ A ist echte Teilmenge von B

3. **Additionssatz**
   $P(A \cup B) = P(A) + P(B) - \mathrm{P}(\mathrm{A} \cap \mathrm{B})$

Aussagenlogik

Distributivität:	**Regeln von de Morgan:**	
$A \wedge (B \vee C) \Leftrightarrow (A \wedge B) \vee (A \wedge C)$	$\neg(A \wedge B) \Leftrightarrow \neg A \vee \neg B$	$(C \setminus A) \cap (C \setminus B) = C \setminus (A \cup B)$
$A \vee (B \wedge C) \Leftrightarrow (A \vee B) \wedge (A \vee C)$	$\neg(A \vee B) \Leftrightarrow \neg A \wedge \neg B$	$(C \setminus A) \cup (C \setminus B) = C \setminus (A \cap B)$

$P(\bar{A} \cap \bar{B}) = P(\overline{A \cup B}) = 1 - P(A \cup B)$

Beispiel Kontraposition

$A \Longrightarrow B \Leftrightarrow \neg A \vee B \Leftrightarrow \neg\neg(\neg A \vee B) \Leftrightarrow \neg(\neg\neg A \wedge \neg B) \Leftrightarrow \neg(A \wedge \neg B) \Leftrightarrow \neg(\neg B \wedge A)$
$\Leftrightarrow \neg\neg B \vee \neg A \Leftrightarrow B \vee \neg A \Leftrightarrow B \Longrightarrow \neg A$

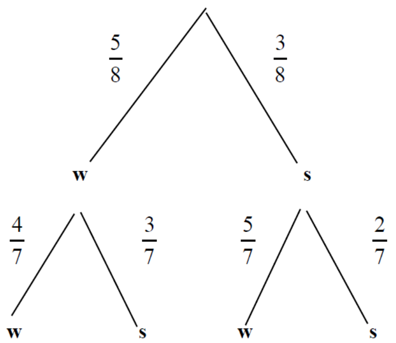
4. Die Wahrscheinlichkeitsdefinition nach Laplace $p = \frac{g}{m}$ ergibt sich ebenfalls aus den Axiomen A1,A2,A3.
   Die Grundereignisse werden als gleichwahrscheinlich vorausgesetzt.
   Wenn $|\Omega| = m$, so hat jedes Grundereignis $\omega \in \Omega$ die Wahrscheinlichkeit $P(\{\omega\}) = \frac{1}{m}$.
   Mit $|A| = g$ ergibt sich $P(A) = g \cdot \frac{1}{m} = \frac{g}{m}$

Mehrstufige Versuche

Beispiel: In einer Urne liegen 5 weisse und 3 schwarze Kugeln.
Man zieht zwei Kugeln nacheinander ohne zurücklegen. Alle Möglichkeiten:

weiss,weiss:	$m = 8 \cdot 7$	$g = 5 \cdot 4$	$P(w,w) = \frac{5\cdot 4}{8\cdot 7}$
weiss,schwarz:	$m = 8 \cdot 7$	$g = 5 \cdot 3$	$P(w,s) = \frac{5\cdot 3}{8\cdot 7}$
schwarz,weiss:	$m = 8 \cdot 7$	$g = 3 \cdot 5$	$P(s,w) = \frac{3\cdot 5}{8\cdot 7}$
schwarz,schwarz:	$m = 8 \cdot 7$	$g = 3 \cdot 2$	$P(s,s) = \frac{3\cdot 2}{8\cdot 7}$

Ereignisbaum



Mehrstufige Versuche kann mal durch einen Ereignisbaum darstellen.

Summer der Äste = 1

Pfadregeln:

- entlang eines Pfades werden die Stufenwahrscheinlichkeiten multipliziert.
- Die Pfade sind disjunkte Ereignisse.
  Ihre Wahrscheinlichkeiten werden addiert.

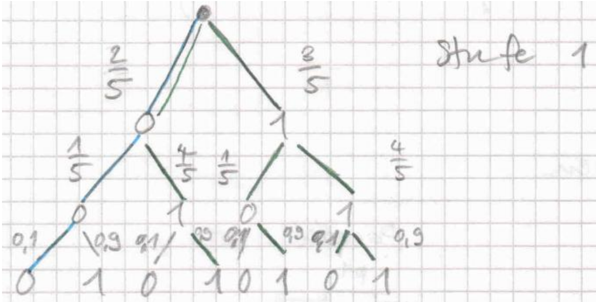
$$P(min.\,1\ schwarze\ Kugel) = P(weiss, schwarz) + P(schwarz, weiss) + P(schwarz, schwarz)$$
$$= \frac{5\cdot 3}{8\cdot 7} + \frac{3\cdot 5}{8\cdot 7} + \frac{3\cdot 2}{8\cdot 7} = \frac{9}{14}$$

Dieselbe Rechnung diesmal aber mit dem Gegenereignis:

$$P(min.\,1\ schwarze\ Kugel) = 1 - P(weiss, weiss) = 1 - \frac{5\cdot 4}{8\cdot 7} = \frac{9}{14}$$

MANIT3 Analysis 3 – Stochastik	HS 2015	Zürcher Hochschule für Angewandte Wissenschaften zh aw School of Engineering
Dozent	Josef Gohl	

Aufgabe 1:

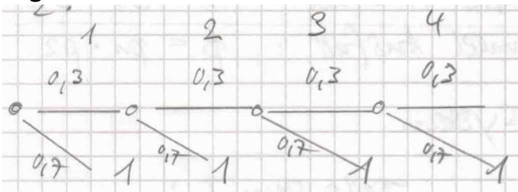


Günstige Pfade:

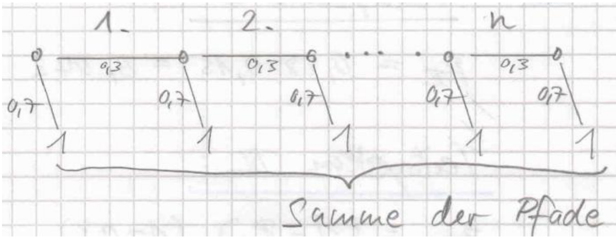
a) $p(0{,}0{,}0) = \frac{2}{5} \cdot \frac{1}{5} \cdot 0.1 = 0.008$

b) $p(min. 2 \text{ "1er"}) = \frac{3}{5} \cdot \frac{4}{5} \cdot 0.9$
$$+\frac{3}{5} \cdot \frac{1}{5} \cdot 0.9 + \frac{3}{5} \cdot \frac{4}{5} \cdot 0.1$$
$$+\frac{2}{5} \cdot \frac{4}{5} \cdot 0.9 =$$

Aufgabe 2:



a) $p(4) = 0.3 \cdot 0.3 \cdot 0.3 \cdot 0.7$
$$= 0.3^3 \cdot 0.7 = 0.0189$$
$$p(n) = 0.7 \cdot 0.3^{(n-1)}$$

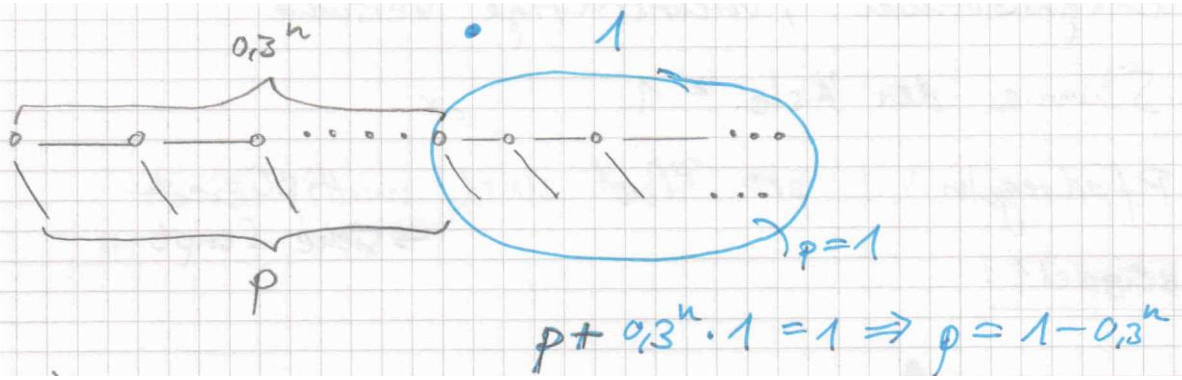


b) $0.7 + 0.7 \cdot 0.3 + 0.7 \cdot 0.3^2$
$$+0.7 \cdot 0.3^{(n-1)}$$

Geometrische Reihe:
$$\frac{1-q^n}{1-q}$$

$$p = 0.7 \cdot \left(1 + 0.3 + \cdots + 0.3^{(n-1)}\right) = 0.7 \cdot \frac{1-0.3^n}{1-0.3} = 0.7 \cdot \frac{1-0.3^n}{0.7} = 1 - 0.3^n$$

Direkt berechnet:



$p(0.9995)$ $\qquad p = 1 - 0.3^n \geq 0.9995$

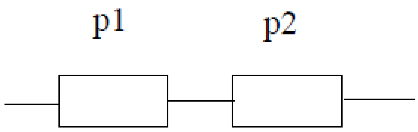
MANIT3 Analysis 3 – Stochastik	HS 2015
Dozent	Josef Gohl

Anwendung 1: Zuverlässigkeit von (komplexen) Systemen

p_i	**Ausfallswahrscheinlichkeit** der i-ten Komponente
$z_i = 1 - p_i$	**Zuverlässigkeit** der i-ten Komponente
P	Ausfallswahrscheinlichkeit des Systems
$Z = 1 - P$	Zuverlässigkeit des Systems

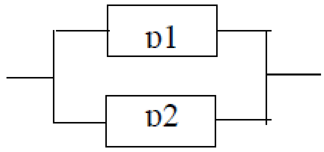
Man stelle sich einen Signalfluss von A nach B vor. Die Komponenten fallen aus (stören den Signalfluss) oder sind intakt. Einfache Systeme bestehen aus zwei Komponenten und diese kann man grundsätzlich auf 2 Arten zusammenschalten:

Serie-Schaltung

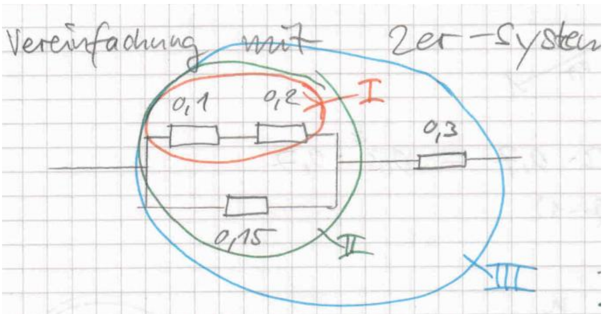


$$P = 1 - z_1 \cdot z_2$$
$$Z = z_1 \cdot z_2$$

Parallel-Schaltung



$$P = p_1 \cdot p_2$$
$$Z = 1 - p_1 \cdot p_2$$



Teilsystem I:
$z_I = (1 - 0.1) \cdot (1 - 0.2) = 0.72$
$p_I = 1 - 0.72 = 0.28$

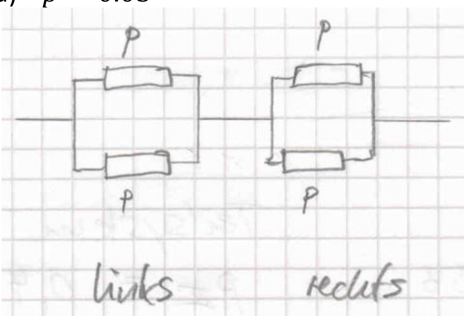
Teilsystem II:
$p_{II} = 0.28 \cdot 0.15 = 0.042$

Teilsystem III:
$z_{III} = (1 - p_{II}) \cdot (1 - 0.3) = 0.958 \cdot 0.7$
$= 0.6706$
$p_{III} = 1 - z_{III} = 1 - 0.6706 = 0.3294$

Beispiel: Ein 4-motoriges Flugzeug kann sich in der Luft halten, wenn

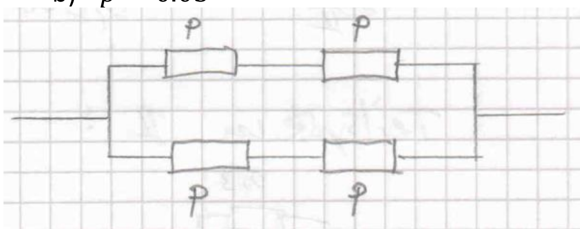
a) Auf jeder Seite mindestens ein Motor intakt ist (Serie-Schaltung)
b) Entweder beide inneren oder beide äusseren Motoren intakt sind (Parallel-Schaltung)

a) $p = 0.05$



$z = z_L \cdot z_R = (1 - p^2)^2$
$z_L = 1 - p^2 = 0.9975$
$z_R = z_L = 0.9975$
$z = z_L \cdot z_R = 0.995$

b) $p = 0.05$



$p_L = 1 - z_{L1} \cdot z_{L2} = 0.0975$
$pR = p_L = 0.0975$
$z = 1 - p_L \cdot p_R = 0.9905$

Anwendung 2: Kryptographische Hashfunktion, Kollisionswahrscheinlichkeit

Man hat N Hashwerte und $n < N$ Versuche. Mit welcher Wahrscheinlichkeit stimmen bei mindestens zwei Versuchen die Hashwerte überein (sogenannte Kollision)? Oder Wie viele Versuche n braucht es, damit die Kollisionswahrscheinlichkeit 0.5 ist?

Anschauliche Demonstration mit dem Geburtstagsproblem: In einer Gruppe von n Personen wettet jemand, dass mindestens 2 Personen den gleichen Geburtstag haben. Wie gross muss n sein, damit die Gewinnchance > 0.5 ist?

Wahrscheinlichkeit für eine Kollision sei P_n. $P_n = 1 - P(alle\ Geburtstage\ verschieden)$
$N = 365$ Tage

$$P_n = 1 - \frac{N}{N} \cdot \frac{N-1}{N} \cdot \frac{N-2}{N} \cdot \ldots \cdot \frac{N-n+1}{N}$$

Kollision $P_3 = 1 - P(alle\ Geburtstage\ verschieden) = 1 - \left(\frac{365}{365} \cdot \frac{364}{365} \cdot \frac{363}{365}\right)$

$$= 1 - 0.991796 = 0.008204$$

$Pn \geq 0.5$: $\quad \frac{N}{N} - \frac{n+1}{N} = \frac{N-n+1}{N}$

Lösung: $\quad n \geq 23$

Allgemein:

Für kleine $|x|$: $\quad e^x \approx 1 + x$

$$1 - \frac{1}{N} = 1 + \frac{-1}{N} \approx e^{-\frac{1}{N}}$$

$$P_n \approx 1 - e^{-\frac{1}{N}} \cdot e^{-\frac{2}{N}} \cdot e^{-\frac{3}{N}} \cdot \ldots \cdot e^{-\frac{n-1}{N}} = 1 - e^{-\frac{1}{2N} \cdot n \cdot (n-1)}$$

$$0.5 \approx 1 - e^{-\frac{1}{2N} \cdot n \cdot (n-1)}$$ nach n auflösen

$$\ln(0.5) \approx -\frac{1}{2N} \cdot n \cdot (n-1)$$

$$\ln(0.5) \approx -\frac{n^2 - n}{2N} \Longrightarrow \ln(2) \approx \frac{n^2 - n}{2N} \approx \frac{n^2}{2N}$$

$$\Longrightarrow n \approx \sqrt{\ln(2) \cdot 2N}$$

Aufgabe : Kollision mit einem bestimmten Hashwert

$$p(min.1) = 1 - (keine\ Kollision) = 1 - \frac{364}{365} \cdot \frac{364}{365} \cdot \ldots \cdot \frac{364}{365}$$

$$= 1 - \left(\left(1 - \frac{1}{N}\right) \cdot \left(1 - \frac{1}{N}\right) \cdot \ldots \cdot \left(1 - \frac{1}{N}\right)\right)$$

$$P_n = 1 - \left(1 - \frac{1}{N}\right)^n = 0.5$$

$$\Longrightarrow n = \frac{\ln(0.5)}{\ln\left(1 - \frac{1}{N}\right)} \approx \frac{\ln(0.5)}{\ln\left(e^{-\frac{1}{N}}\right)} = \frac{\ln(0.5)}{-\frac{1}{N}} = N \cdot \ln(2)$$

$$\boxed{\approx e^{\frac{1}{N}}}$$

Bedingte Wahrscheinlichkeit

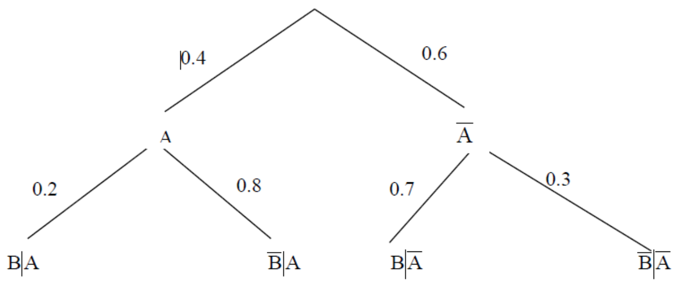
Verhältnis von Wahrscheinlichkeiten. Die Bedingung steht im Nenner.

$P(A|B)$: Wahrscheinlichkeit für A unter der Bedingung B.

$$P(A|B) = \frac{P(A \cap B)}{P(B)} \leq 1$$

$$\left.\begin{aligned} P(A|B) &= \frac{P(A \cap B)}{P(B)} \\ P(B|A) &= \frac{P(B \cap A)}{P(A)} \end{aligned}\right\} \Longrightarrow P(A \cap B) = P(B) \cdot P(A|B) = P(A) \cdot P(B|A)$$

Siehe **Ereignisbaum** → Multiplikation entlang eines Pfades.



Totale Wahrscheinlichkeit:

$$P(B) = P(B|A) \cdot P(A) + P(B|\bar{A}) \cdot P(\bar{A})$$

Beispiel: $P(B) = 0.2 \cdot 0.4 + 0.7 \cdot 0.6 = 0.5$

Satz von Bayes:

$$P(A|B) = \frac{P(A \cap B)}{P(B)} = \frac{P(B|A) \cdot P(A)}{P(B|A) \cdot P(A) + P(B|\bar{A}) \cdot P(\bar{A})}$$

Verteilungen

Zufallsvariable

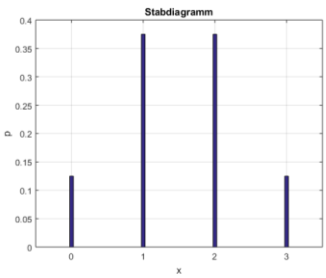
In statistischen Fragen werden die Ereignisse als Zahlvariable X defineirt. Ordnet man jedem Ereignis einen Zahlenwert zu, so nimmt die Variable X diese Werte zufällig an. Z.B. $X =$ Augensumme beim Werfen von zwei Würfeln. X ist eine **Zufallsvariable** und die Wahrscheinlichkeiten sind als $P(X)$ auszudrücken.

Wahrscheinlichkeitsverteilung von $\boldsymbol{X}$: $X \longrightarrow P(X)$

Beispiel 1:
Man wirft eine Münze (Kopf, Zahl) drei Mal. X sei die Anzahl ‚Kopf'. X nimmt die Werte 0,1,2 oder 3 an. Zugehörige Wahrscheinlichkeiten:
z.B. $X = 2$: $P(X=2) = \frac{g}{m} = \frac{3}{2^3} = \frac{3}{8} = 0.375$

Wahrscheinlichkeiten werden als Tabelle oder graphische als Stabdiagramm dargestellt.

$\boldsymbol{X}$	$\boldsymbol{P(X)}$
0	0.125
1	0.375
2	0.375
3	0.125



Diskrete und stetige Zufallsvariablen

Diskret	Stetig
- Endlich viele Werte - Stabdiagramm - Stabhöhe = Wahrscheinlichkeit - Summe der Wahrscheinlichkeiten = 1 - Zufallsvariable als Zählvariable X zählt z.B. die Anzahl Erfolge bei einem Zufallsexperiment	- Alle Werte eines Intervalls aus $\mathbb{R}$ - Eine stetige Zufallsvariable misst eine Grösse → Messdaten **Wahrscheinlichkeits-Dichtefunktion $\boldsymbol{f(x)}$**  $f(x) \geq 0$ $f(x)$ definiert für $x \in \mathbb{R}$ Die gesamte Fläche unter der Dichtekurve ist $\mathbf{1}$

Verteilungsfunktion

Grundlegend für jede Zufallsvariable (diskret oder stetig). Berechnete **kumulierte** Wahrscheinlichkeiten P_{kum}.

Die Verteilungsfunktion ist der grundlegende Begriff zur Beschreibung der Wahrscheinlichkeits-verteilung.

(kumulierte) **Verteilungsfunktion** der Zufallsvariablen X:

$$F(X) = P(X \leq x) \qquad x \in \mathbb{R}$$

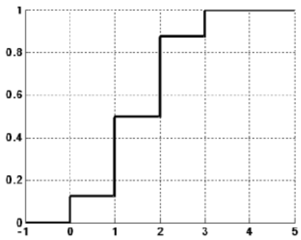
$F(x)$ ist für jedes reelle x definiert.

Für diskrete Zufallsvariable X gilt:

$$F(X) = \sum_{x_j \leq x} p(x_j)$$

$\boldsymbol{X}$	$\boldsymbol{P(X)}$	$\boldsymbol{P_{kum} = F(x)}$
0	0.125	0.125
1	0.375	0.5
2	0.375	0.875
3	0.125	1

$F(0) = 0.125$
$F(0.999) = 0.125$
$F(1) = 0.5$
$F(3) = 1.000$
$F(5) = 1.000$
$F(-2) = 0$



$F(x)$ ist eine monoton wachsende Funktion. Die Wahrscheinlichkeiten nehmen sprunghaft zu.

Hypergeometrische Verteilung

Binomialverteilung

Beschreibt die Wiederholung eines elementaren Versuchs. Anzahl Schritte n vorgegeben. Man führt den Versuch n mal durch. Es kann nur zwei Ergebnisse haben: Erfolg oder Misserfolg.

	Indikatorvariable I	P
Erfolg	1	p
Misserfolg	0	$1-p$

p: Erfolg

Beispiel: $n = 7, k = 3$

Die Sequenz 0001011 hat die Wahrscheinlichkeit $p^3 \cdot (1-p)^4$. $1 = p, \quad 0 = 1 - p$

Es gibt $\frac{7!}{3! \cdot 4!} = \binom{7}{3}$ günstige Sequenzen. Alle haben diese Wahrscheinlichkeit.
Somit ist

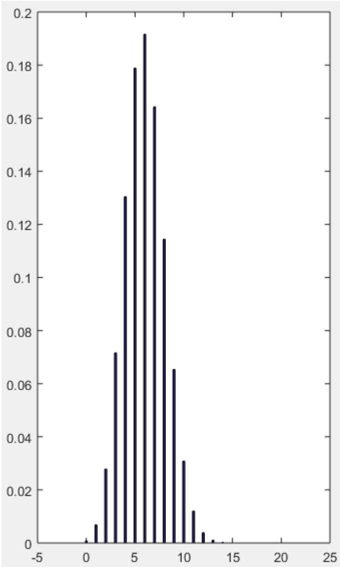
$$P(X = 3) = \binom{7}{3} \cdot p^3 \cdot (1-p)^4$$

Allgemein: Führt man einen Versuch mit der Erfolgswahrscheinlichkeit p n mal durch, so ist die Wahrscheinlichkeit für $X = k$ Erfolge.

$$P(X = K) = \binom{n}{k} \cdot p^k \cdot (1-p)^{n-k} \qquad k = 0,2, \dots, n$$

Die Zufallsvariable X ist **binomial** verteilt:
$X \sim B(p, n)$

MATLAB Beispiel: $X \sim B(0.3,20)$



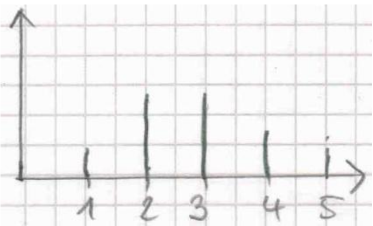

```
x = 0:20;
y = binopdf(x,20,0.3);
bar(x,y);
```

Die Wahrscheinlichkeit ist am grössten (Modus) bei $x = 6$.

$6 = 0.3 \cdot 20 = p \cdot n$

D.h. die Wahrscheinlichkeiten sind gross in der Nähe von $n \cdot p$.

Beispiel 8:
5% der Maroni eines Händlers sind schlecht. Jemand kauft 50 Maroni. X ist die Anzahl schlechte Maroni. Modell: $X \sim B(0.05,50)$ grösste Wahrscheinlichkeit $p \cdot n = 0.05 \cdot 50 = 2.5$



a) Wahrscheinlichkeit, dass höchstens 4 schlecht sind:

$$P(X = k) = P(X \leq 4) = \binom{50}{k} \cdot 0.05^k \cdot 0.95^{50-k}$$
$$= \binom{50}{0} \cdot 0.05^0 \cdot 0.95^{50} + \binom{50}{1} \cdot 0.05^1 \cdot 0.95^{49} + \cdots + \binom{50}{4} \cdot 0.05^4 \cdot 0.95^{46} = 0.8964 = F(4)$$

b) Wahrscheinlichkeit, dass mindestens eine schlecht ist:

$$P(X \geq 1) = 1 - P(X \leq 0) = 1 - \binom{50}{0} \cdot 0.05^0 \cdot 0.95^{50} = 1 - 0.95^{50} = 0.9231$$

c) Wie viele Maroni müsste man kaufen, damit mit Wahrscheinlichkeit 0.99 mindestens eine schlechte dabei ist?

$X \sim B(0.05, n)$

$P(X \geq 1) = 1 - 0.95^n = 0.99$ Auflösen nach n mit Solver → $n = 90$

Geometrische Verteilung

Warten (Versuche) bis der Erfolg eintritt. X ist die Anzahl Misserfolge (Warteschritt), bis der erste Erfolg eintritt. $p =$ Misserfolgswahrscheinlichkeit.

Beispiel: $x = 5$ Sequenz: $11111\mathbf{0}$

$$F(x) = P(X \leq k) = (1-p) \cdot [\underbrace{1 + p + p^2 + \cdots + p^k}_{\text{geometrische Reihe}}]$$

$$= (1-p) \cdot \frac{1-p^{k+1}}{1-p} = 1 - p^{k+1} = F(k)$$

$$P(X = K) = p^k \cdot (1-p)$$

Die Zufallsvariable X ist **geometrisch** verteilt:
$X \sim GM(p)$

Die Wahrscheinlichkeiten entsprechen der Exponentialfunktion.

Beispiel 9:
Man setzt beim Roulette auf Rot. X ist die Anzahl Misserfolge bis Rot erscheint. Bestimme die Verteilung von X. $X \sim GM\left(p = \frac{19}{37}\right)$.

a) Mit welcher Wahrscheinlichkeit erscheint genau beim 3. Versuch zum ersten mal ROT?
$$P(X = 2) = \frac{19-1}{37} \cdot \left(\frac{19}{37}\right)^2 = \frac{18}{37} \cdot \left(\frac{19}{37}\right)^2 = 0.1283$$

b) Mit welcher Wahrscheinlichkeit erscheint spätestens beim 4. Versuch zum ersten mal ROT?
$$P(X \leq 3) = F(3) = 1 - p^4 = 1 - \left(\frac{19}{37}\right)^4 = 0.93$$

c) Mit welcher Wahrscheinlichkeit erscheint frühestens beim 5. Versuch zum ersten mal ROT?
$P(X \geq 4) = 1 - P(X \leq 3) = 1 - F(3) = 1 - (1 - p^4) = p^4 = 1 - 0.93 = 0.07$

Vergleich der Binomialverteilung mit der Hypergeometrischen

Da die binomialen Wahrscheinlichkeiten einfacher zu berechnen sind, kann man die hypergeometrische Verteilung approximativ durch die binomiale Verteilung berechnen.

Beispiel 10:
$x = 3$ $\qquad P(X = 3) = \binom{15}{3} \cdot 0.4^3 \cdot 0.6^{12} = 0.063$ $\qquad$ (Binomialverteilung)

$x = 3$ $\quad P(X = 3) = \frac{\binom{20}{3}\cdot\binom{30}{12}}{\binom{50}{15}} = 0.044$ (Hypergeometrische Verteilung)

Poisson-Verteilung

Bekannt ist der Mittelwert der Verteilung. Gleiche mittlere Anzahl Treffer λ.
Innerhalb von n Intervallen I der Länge l hat es im Mittel $N = \lambda \cdot n$ Treffer.
D.h. die Anzahl Treffer in I ist binomial verteilt: $X \sim B(\frac{1}{n}, N)$

Für k Treffer in I gilt die Wahrscheinlichkeit $P(X = k) = \binom{n}{k} \cdot \left(\frac{1}{n}\right)^k \cdot \left(1 - \frac{1}{n}\right)^{N-k}$

Der Erwartungswert für jedes n ist demnach $\quad \mu = \frac{1}{n} \cdot N = \lambda$.

Grenzmodell

$n \to \infty$

Demonstration mit $k = 4$.

$$\frac{N}{n} = 4 = \lambda \Longrightarrow \frac{1}{n} = \frac{4}{N} \Longrightarrow n = \frac{N}{4}$$

Die Wahrscheinlichkeiten stabilisieren sich für $n \to \infty$.

$$P(X = k) = \binom{N}{k} \cdot \left(\frac{1}{n}\right)^k \cdot \left(1 - \frac{1}{n}\right)^{N-k} \xrightarrow{n\to\infty} e^{-\lambda} \cdot \frac{\lambda^k}{k!}$$

Die Varianz hat den Grenzwert $n \cdot p \cdot (1 - n) = 4 \cdot (1 - p) = 4 \Longrightarrow p$ strebt gegen 0

Poisson-Verteilung mit dem Paramter $\lambda = \mu = \sigma^2$: $\quad X \sim Po(\lambda)$

Für grosse $n$und kleine p $(n \geq 50,\ p \leq 0.05)$ und wenn $n \cdot p$ nicht zu gross ist, ist die Poisson-Verteilung eine gute Näherung für die Binomialverteilung $X \sim B(p, n)$ mit $\lambda = n \cdot p$.

$$X \sim B(p, n) \approx Po(n \cdot p)$$

Aus kleinen Trefferwahrscheinlichkeiten ergibt sich die Interpretation, dass die Poisson-Verteilung seltene Ereignisse gut beschreibt. (Waldbrände, Schadenfälle).

Beispiele :

$\lambda = 4$, Erwartungswert ist 4. Ist die Summe = 1 ?

$$e^{-4} \cdot \frac{4^4}{4!} = 0.195367$$

$$\sum_{i=0}^{\infty} p_i = 1$$

$$e^{-\lambda} \cdot \sum_{k=0}^{\infty} \frac{\lambda^k}{k!} = e^{-\lambda} \cdot e^{\lambda} = 1$$

Reihe für e^{λ}

Beispiel 1 : $\lambda = 4$

a) $P(X = 3)$ $\quad e^{-4} \cdot \frac{4^3}{3!} = 0.195367$

b) $P(X \leq 4)$ $\quad e^{-4} \cdot \left(\frac{4^0}{0\,!} + \frac{4^1}{1\,!} + \frac{4^2}{2\,!} + \frac{4^3}{3\,!} + \frac{4^4}{4\,!}\right) = e^{-4} \cdot \left(1 + \lambda + 8 + \frac{20}{3} + \frac{20}{3}\right) = 0.482312$

c) $P(X \geq 2) = 1 - P(X \leq 1) = 1 - e^{-4} \cdot (1 + 4) = 0.908422$

Beispiel 2 :

a) $\mu = \frac{45}{60} \cdot 2 = 1.5 = \lambda$
$P(X \leq 1) = e^{-1.5} \cdot (1 + \lambda) = 0.557825$

b) $\lambda = \frac{45}{60} \cdot 3 = \frac{9}{4} = 2.25$
$P(X < 3) = 1 - P(X \leq 3) = 1 - poisscdf(3{,}2.25) = 0.190567$

Schätzen

Anwendung des zentralen Grenzwertsatzes
Geschätzte Grössen symbolisiert man mit ^.

Punktschätzung

Intuitiv wird der Durchschnitt (= Schätzfunktion) berechnet.

Parameter: Mittelwert μ und Anteil p

$$\hat{\mu} = \hat{x}$$
$$\hat{p} = \frac{A}{n} = a$$

$\hat{p}$ ist der relative Anteil der Anzahl Erfolge.

Beispiel:
Stichprobe von 120 Marroni (Gewichte in Gramm)

- Mittleres Gewicht aller Marroni des Händlers geschätzt: $\hat{\mu} = \hat{x} = 19.37$
- Anteil wurmstichiger Marroni aller Marroni des Händlers $\hat{p} = a = \frac{32}{120} = 0.267$

Intervallschätzung

Eine Angabe für die **Genauigkeit der Punktschätzung**.

Man verwendet den zentralen Grenzwertsatz, für grosse Stichproben $(n \geq 100)$ gilt:

$$T = \frac{\bar{X} - \mu}{\frac{s}{\sqrt{n}}} \approx N(0,1)$$

Stichprobenvarianz

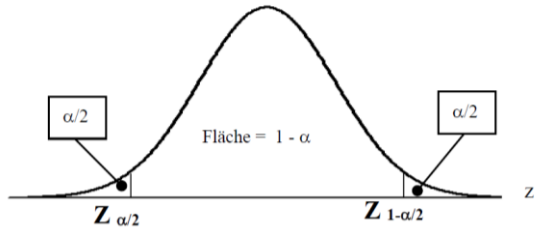
$$\boldsymbol{s^2} = \frac{1}{n-1}\sum_{i=1}^{n}(x_i - \bar{x})^2 = \frac{1}{n-1}\left[\sum_{i=1}^{n} x_i^2 - n \cdot \bar{x}^2\right]$$

Standardfehler der Stichprobe

$$\boldsymbol{\hat{\sigma}_{\bar{x}}} = \frac{s}{\sqrt{n}}$$

Vertrauensintervall

Die Zufallsvariable T liegt mit Wahrscheinlichkeit $1-\alpha$ zwischen den Quantilen der $N(0,1)$ Verteilung $\quad Z_{\frac{\alpha}{2}} = -Z_{1-\frac{\alpha}{2}}$ und $Z_{1-\frac{\alpha}{2}}$



$$-Z_{1-\frac{\alpha}{2}} \leq \frac{\bar{X}-\mu}{\frac{s}{\sqrt{n}}} \leq Z_{1-\frac{\alpha}{2}}$$

Durch Umformung erhält man:

$$\bar{X} - \frac{s}{\sqrt{n}} \cdot Z_{1-\frac{\alpha}{2}} \leq \mu \leq \bar{X} + \frac{s}{\sqrt{n}} \cdot Z_{1-\frac{\alpha}{2}}$$

Da $\bar{X}$ eine Zufallsvariable ist, so ist auch das Intervall $\bar{X} \pm \frac{s}{\sqrt{n}} \cdot z_{1-\frac{\alpha}{2}}$ variabel.

Es überdeckt mit Wahrscheinlichkeit $1-\alpha$ den festen aber unbekannten Wert μ.

$1-\alpha$ **Vertrauensintervall** für $\mu \qquad (1-\alpha\, VI)$

$$\bar{x} \pm e = [\bar{x}-e, \bar{x}+e]$$

mit $\quad e = \frac{s}{\sqrt{n}} \cdot Z_{1-\frac{\alpha}{2}}$