Münster University of Applied Sciences

Verteilungen & Bestimmungen

Inhalt:

- Bernoulli und die zwei Möglichkeiten
- Von der Datenmenge zur Erkenntnis
- Doch ganz dicht! Kontinuierliche Verteilungen
- In vielen Dimensionen: Fehlerfortpflanzung, Korrelationen etc.

Prof. Dr. Martin Poppe

Bernoulli und die 2 Möglichkeiten

Kopf oder Zahl?
Sein oder nicht sein?
Im Tor oder vorbei?

Wenn 90 % der Strafstöße verwandelt werden, wie groß ist die Wahrscheinlichkeit, dass von 20 genau 17 drin sind?

Jacob Bernoulli
(1655 -1705)

Wenn ich *n* mal mit der Wahrscheinlichkeit *p* teste, wie groß ist dann die Wahrscheinlichkeit, genau *k* gute Ergebnisse zu bekommen?

Bernoulli und die 2 Möglichkeiten

Ergebnisenge für 3 Würfe:

$\omega =$ (000), (001), (010), (011),
(100), (101), (110), (111)

Wahrscheinlichkeiten, sortiert:

0 Köpfe $P(000) = (1-p)^3$

1 Kopf $P(001) = P(010) = P(100) = p \cdot (1-p)^2$

2 Köpfe $P(110) = P(101) = P(011) = p^2 \cdot (1-p)$

3 Köpfe $P(111) = p^3$

Bernoulli und die 2 Möglichkeiten

Ergebnisenge für 3 Würfe:

$\omega =$ (000), (001), (010), (011), (100), (101), (110), (111)

Wahrscheinlichkeiten, summiert:

0 Köpfe $P(k=0)=(1-p)^3$

1 Kopf $P(k=1)=3\cdot p\cdot(1-p)^2$

2 Köpfe $P(k=2)=3\cdot p^2\cdot(1-p)$

3 Köpfe $P(k=3)=p^3$

Bernoulli und die 2 Möglichkeiten

Verallgemeinerung 3 → n

Jedes Elementarereignis
mit *k* Köpfen hat
die Wahrscheinlichkeit

$$P_i(k) = p^k \cdot (1-p)^{(n-k)}$$

Aber zu jeder Anzahl von *k* gibt
es viele Elementareregnisse:

$$\binom{n}{k} = \frac{n \cdot (n-1) \cdot \ldots (n-k+1)}{1 \cdot 2 \cdot \ldots k}$$

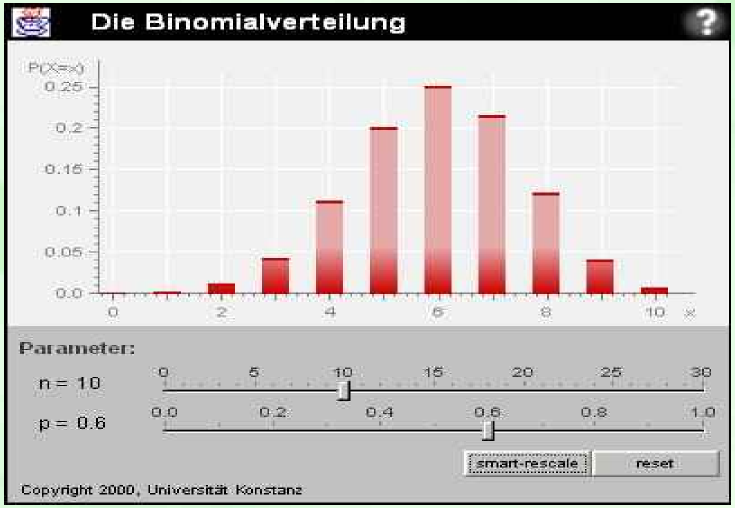
$$\binom{n}{k} = \frac{n!}{k! \cdot (n-k)!}$$

Bernoulli und die 2 Möglichkeiten

Binomialverteilung

Alle Elementarereignisse mit der gleichen Kopfzahl:

$$P(n,k)=\binom{n}{k}\cdot p^k(1-p)^{(n-k)}$$



(Applet der Uni Konstanz)

Bernoulli und die 2 Möglichkeiten

Binomialverteilung

Relevanz der allgemeinen Formel

$$P(n,k)=\binom{n}{k}\cdot p^k(1-p)^{(n-k)}$$

Zu jeder ja/nein Entscheidung Verteilungsfunktion bekannt!

Mächtiges Werkzeug zur Bestimmung und Eingrenzung von Wahrscheinlichkeiten

Bernoulli und die 2 Möglichkeiten

Derivate der Binomialverteilung

Original **Binomialverteilung**:
Wahrscheinlichkeit für *k* bei gegebenem *n*

$$P_{original}(n,k)=\binom{n}{k}\cdot p^k(1-p)^{(n-k)}$$

negative Binomialverteilung
Wahrscheinlichkeit für *n* bei gegebenem *k*

$$P_{negativ}(n,k)=\binom{n-1}{k-1}\cdot p^k(1-p)^{(n-k)}$$

Bernoulli und die 2 Möglichkeiten

Derivate der Binomialverteilung

„Wie groß ist die Wahrscheinlichkeit, 3 Hauptgewinne im Lotto mit genau 7 abgegbenen Scheinen zu bekommen?“

negative Binomialverteilung
Wahrscheinlichkeit für *n* bei gegebenem *k*

$$P_{negativ}(n,k)=\binom{n-1}{k-1}\cdot p^{k}(1-p)^{(n-k)}$$

Normierung (Summe über *n*!): $\sum_{n=1}^{\infty} P(n,k)\to 1$

Vorsicht: subtile Formulierungsunterschiede!... denn

Bernoulli und die 2 Möglichkeiten

Derivate der Binomialverteilung

„Wie groß ist die Wahrscheinlichkeit, mit 7 abgegebenen Scheinen genau 3 Hauptgewinne im Lotto zu bekommen?"

ergibt die Original- Binomialverteilung
Wahrscheinlichkeit für *k* bei gegebenem *k*

$$P_{original}(n,k)=\binom{n}{k}\cdot p^k(1-p)^{(n-k)}$$

Normierung (Summe über *k*!): $\sum_{k=0}^{n} P(n,k)\rightarrow 1$

Präzise Formulierungen durch nichts zu ersetzen!

Bernoulli und die 2 Möglichkeiten

Derivate der Binomialverteilung

Bitte formulieren Sie 3 Fragen so, dass ihre Beantwortung mit Hilfe der negativen Binomialverteilung berechnet werden kann.

Bitte formulieren Sie dann 3 Fragen, die mit der Original-Binomialverteilung beantwortbar sind.

Bernoulli und die 2 Möglichkeiten

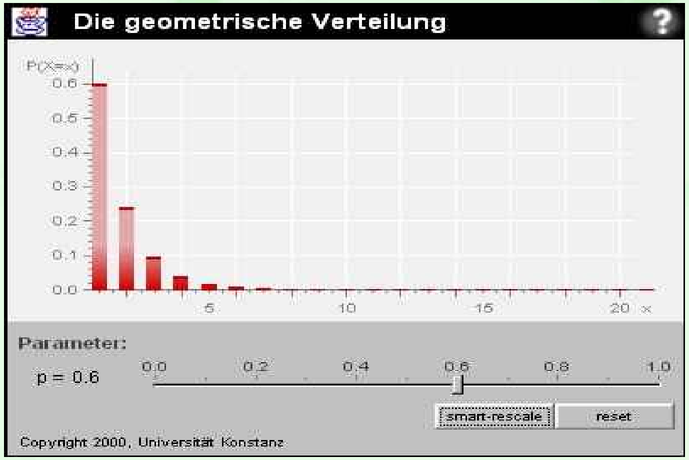
Derivate der Binomialverteilung

Wie viele Versuche bis zum ersten günstigen Ereignis?

Exp. Nr.	1	2	3	…	k	k+1	k+2	…	n
Ergebnis	0	0	0	…	1	egal	egal	egal	egal

Geometrische Verteilung
Wahrscheinlichkeit, dass *k* das erste günstige Ereignis

$$P_{geo}(k = Erstwert) = (1-p)^{(k-1)} \cdot p$$



(Applet der Uni Konstanz)

Bernoulli und die 2 Möglichkeiten

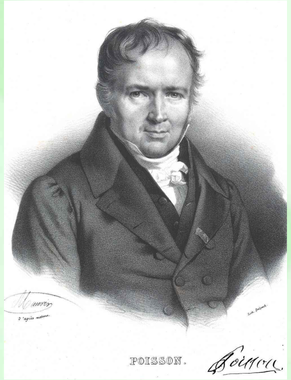
Derivate der Binomialverteilung

Spezialfall: Binomialverteilung für große n und sehr kleine p mit $\mu = n \cdot p$

Poisson (1837) „*Recherches sur la probabilité des jugements en matières criminelles et en matière civile*"

Poisson Verteilung

$$P(\mu, k) = \frac{\mu^k}{k!} e^{-\mu}$$



Siméon Denis Poisson

Welche Bedeutung hat μ?

Was ist egal?

Bernoulli und die 2 Möglichkeiten

Zusammenfassung

„*P(k)* bei gegebenem *n*“ → Bernoulli
Speziell für n groß, p klein → Poisson

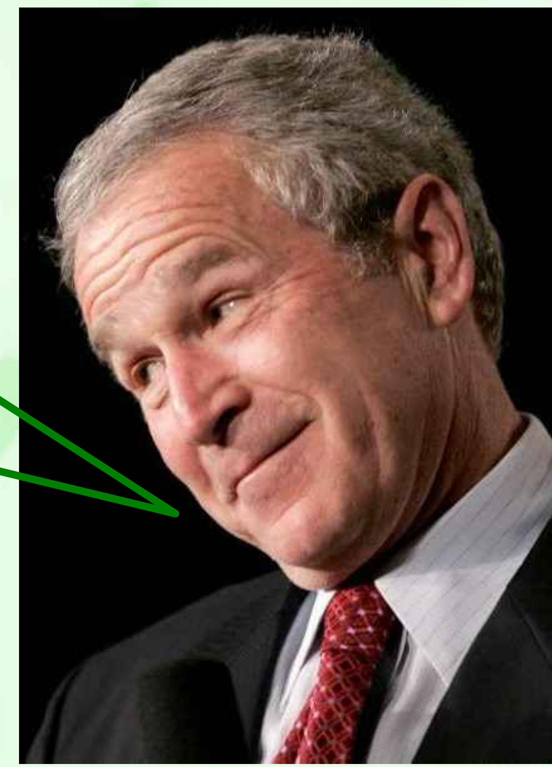
„*P(n)* bei gegebenem *k*“ → negativ Bernoulli

„*P(k=erstes günstiges)*“ → geometrische

Daten → Erkenntnis

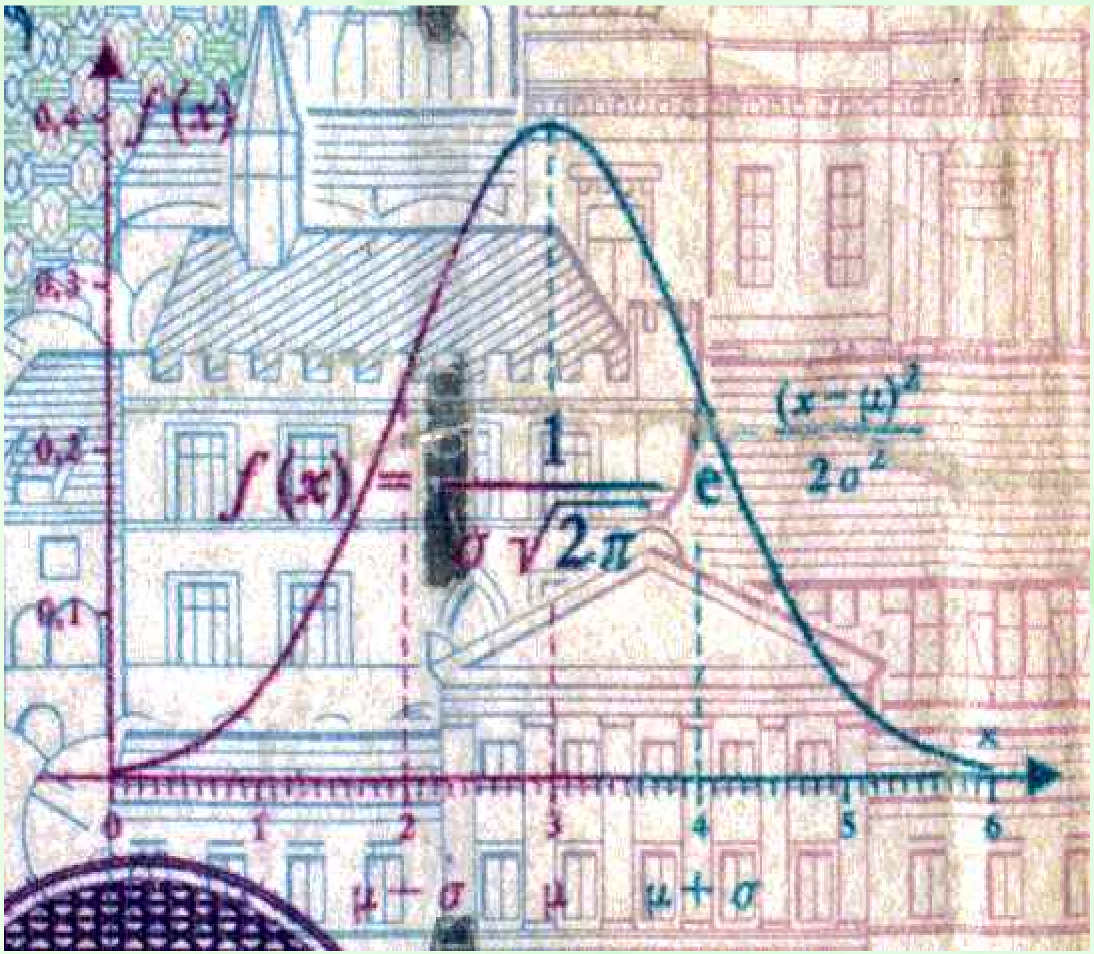
Parameter bestimmen *ist auch nichts anderes als* ***Hypothesen testen!***

Das musst Du mir erklären!



Daten → Erkenntnis

Nützliche Definitionen



Erwartungswert

$$\mu = E(x) = \sum_i p_i x_i$$

$$\mu_{exp.} = \frac{1}{n} \sum_{i=0}^{n} x_i$$

Varianz

$$V(x) = \sum_i p_i \cdot (x_i - \mu)^2$$

$$V_{exp} = \frac{1}{n-1} \sum_{i=0}^{n} (\mu - x_i)^2$$

Standard Abweichung

$$\sigma_x = \sqrt{V(x)}$$

Daten → Erkenntnis

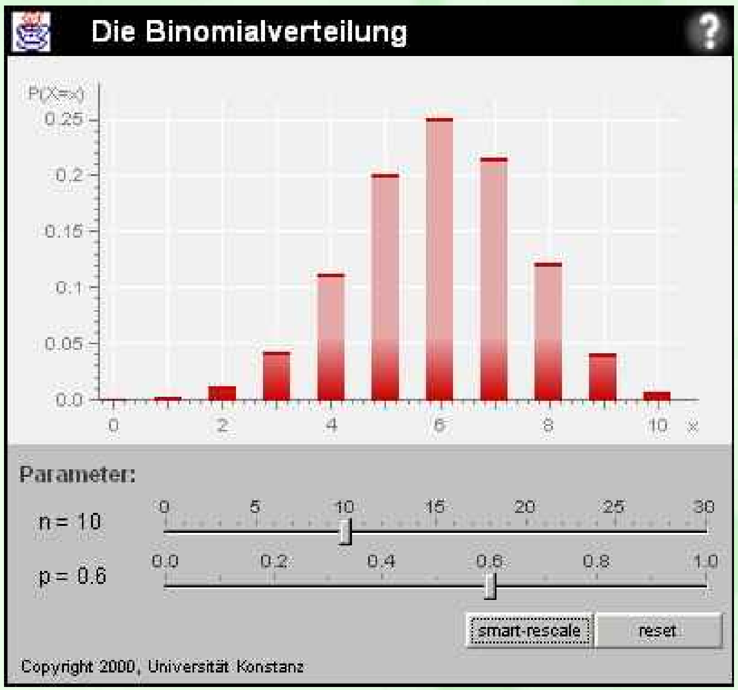
Verschiebungssatz der Varianz

$$V(x)=\sum_i p_i\cdot(x_i-\mu)^2$$
$$V(x)=\sum_i p_i\cdot(x_i^2-2\mu x_i+\mu^2)$$
$$V(x)=\sum_i p_i\cdot x_i^2-2\mu\sum_i p_i\cdot x_i+\mu^2\sum_i p_i$$
$$V(x)=E(x^2)-2\mu\cdot E(x)+\mu^2$$
$$V(x)=E(x^2)-(E(x))^2$$
$$V(x)=E(x^2)-\mu^2$$

„Erwartungswert des Quadrates minus Quadrat des Erwartungswertes“

Daten → Erkenntnis

Eigenschaften der Binomialfunktion:



(Applet der Uni Konstanz)

$$P(n,k)=\binom{n}{k}\cdot p^k(1-p)^{(n-k)}$$

$$E(k)=\mu=n\cdot p$$

$$V(k)=\sigma_k^2=n\cdot p\cdot(1-p)$$

$\rightarrow$

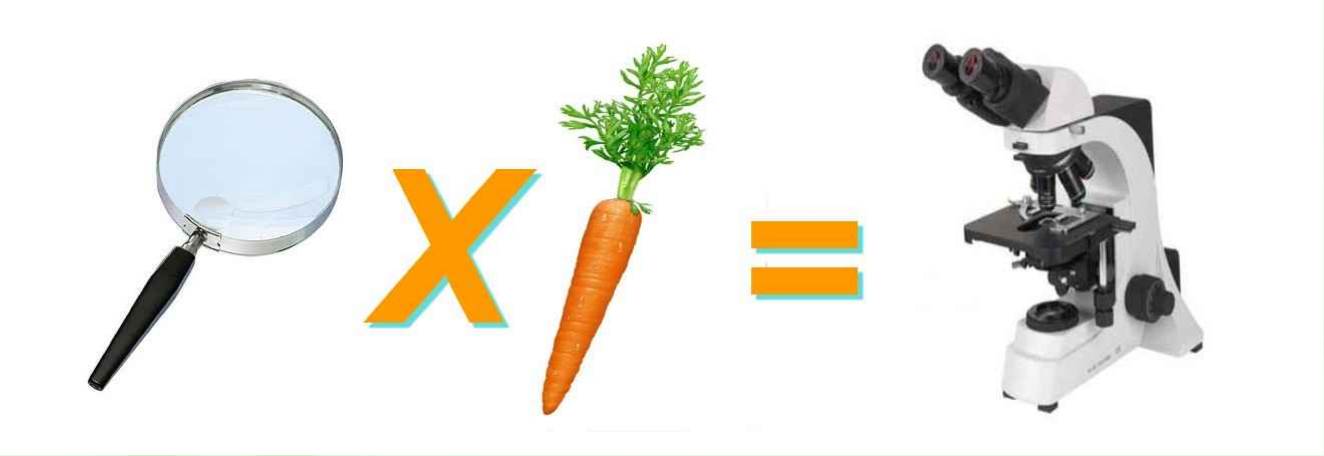
$$\frac{\sigma_k}{\mu}=\frac{\sqrt{n\cdot p\cdot(1-p)}}{np}=\sqrt{\frac{1-p}{p}}\cdot\sqrt{\frac{1}{n}}$$

Relative Streubreite sinkt mit der Wurzel der Testanzahl!

Daten → Erkenntnis

Eigenschaften der Binomialfunktion:
Das Wurzelgesetz

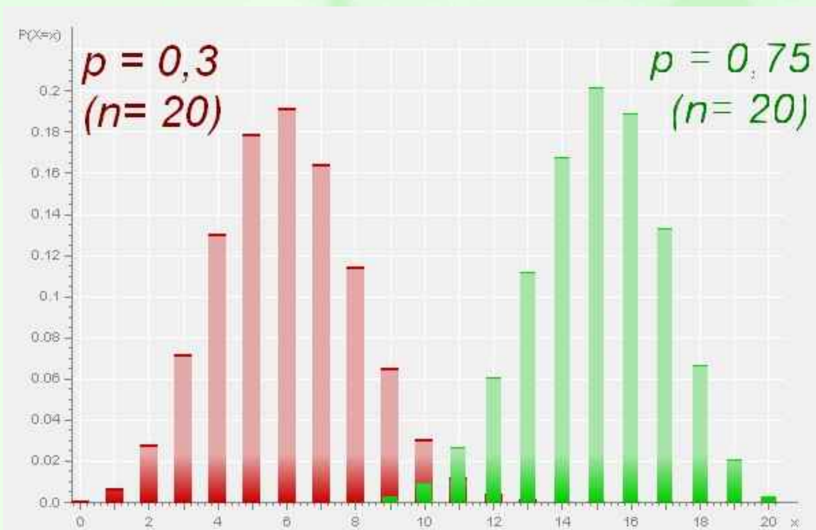
$$P(n,k)=\binom{n}{k}\cdot p^k(1-p)^{(n-k)}$$



$$\frac{\sigma_k}{\mu}=\frac{\sqrt{n\cdot p\cdot(1-p)}}{np}=\sqrt{\frac{1-p}{p}}\cdot\sqrt{\frac{1}{n}}$$

Daten → Erkenntnis

Eigenschaften der Binomialfunktion:
Experimentelle Bestimmung von *p*



Wenn $p = 0,3$ Wenn $p = 0,75$

$P(k > 9) \approx 0,05$ $P(k < 11) \approx 0,02$

Was bedeutet $k_{exp} = 10$

Bei $n = 20$?

95% Wahrscheinlichkeit: $p < 0,3$ ausgeschlossen

98% Wahrscheinlichkeit: $P > 0,75$ ausgeschlossen.

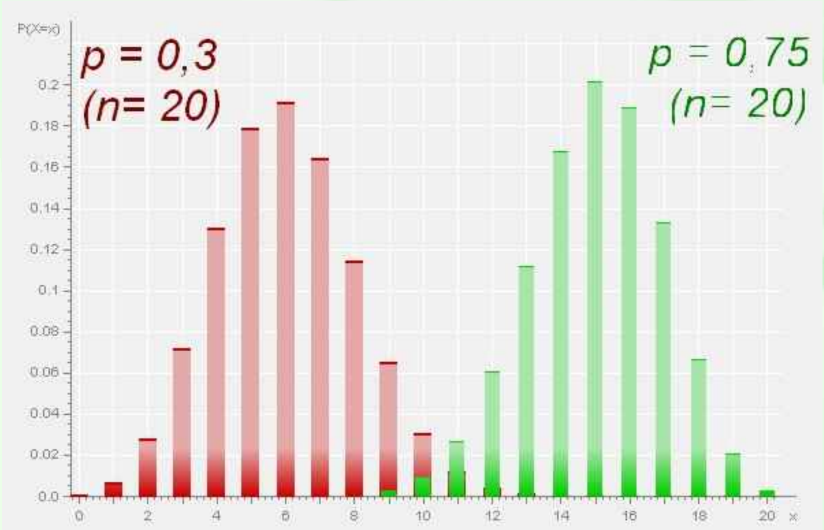
→

93% Wahrscheinlichkeit:

$0,3 < p < 0,75$

Daten → Erkenntnis

Eigenschaften der Binomialfunktion:
Bestimmung des Vertrauensintervalls



Das „y% Vertrauensintervall“ gibt an mit welcher y%-igen Wahrscheinlichkeit p in diesem Intervall liegt.

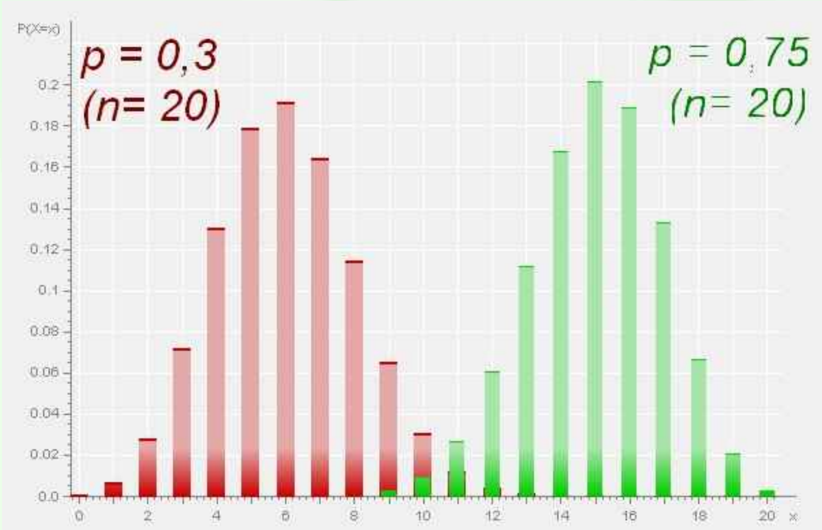
Beispiel 90 %, numerisch lösen:

$$\sum_{k=0}^{k_{exp.}} P(n,k) = \binom{n}{k} \cdot p_{min}^{k} (1 - p_{min})^{(n-k)} > 0{,}95$$

$$\sum_{k=0}^{k_{exp.}} P(n,k) = \binom{n}{k} \cdot p_{max}^{k} (1 - p_{max})^{(n-k)} < 0{,}05$$

Daten → Erkenntnis

Verallgemeinerung des Verfahrens zur Bestimmung des Vertrauensintervalls



Wenn die Verteilungsfunktion *p(x)* für eine Messgröße *x* bekannt ist, dann ist die Wahrscheinlichkeit *y*, dass der Erwartungswert im Bereich $[x_{min}, x_{max}]$ liegt

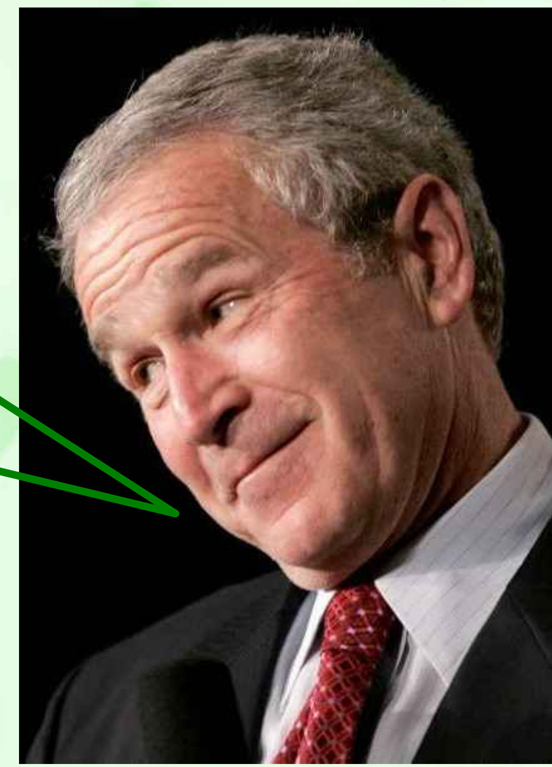
$$\sum_{x_i=-\infty}^{x_{exp.}} P(x_i, x_{min}) > 1 - y/2$$

$$\sum_{x_i=-\infty}^{x_{exp.}} P(x_i, x_{max}) < y/2$$

Daten → Erkenntnis

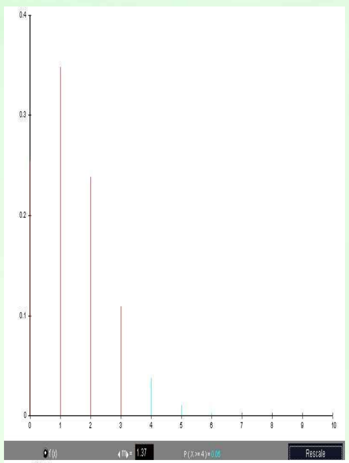
*Die **Grenzen des Vertrauensintervalls** markieren die Punkte, jenseits derer die **Theorien mit dem Experiment unverträglich** werden*

Education mission accomplished!



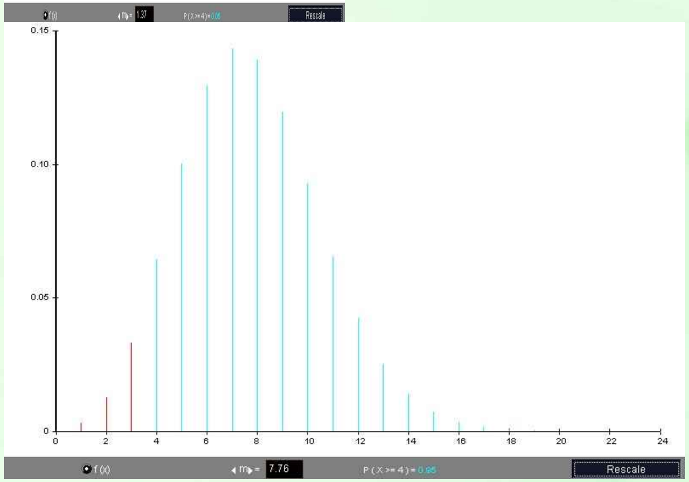
Daten → Erkenntnis

Spezialfall Poisson-Verteilung:
Experimentelle Bestimmung von μ
Bestimmung des 90% Vertrauensintervalls für $k_{exp} = 3$:



$$P(\mu = 1{,}37\,, k > 3) = 0{,}05$$

$$P(\mu = 7{,}76\,, k > 3) = 0{,}95$$



(www.stat.wvu.edu/SRS/Modules/Poisson/poisson.html)

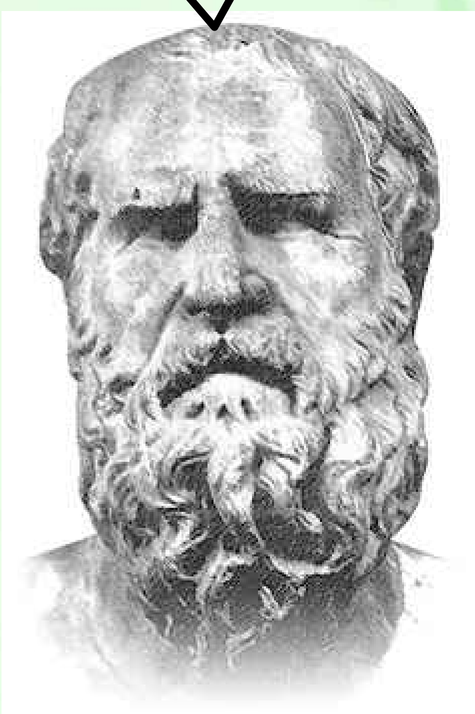
Mit 90% Wahrscheinlichkeit gilt

$$\mu = 3^{+4{,}76}_{-1{,}63}$$

(PS: meist wird nur das 68% Intervall angegeben)

Kontinuierliche Verteilungen

πάντα ῥεῖ,
„alles fließt“)

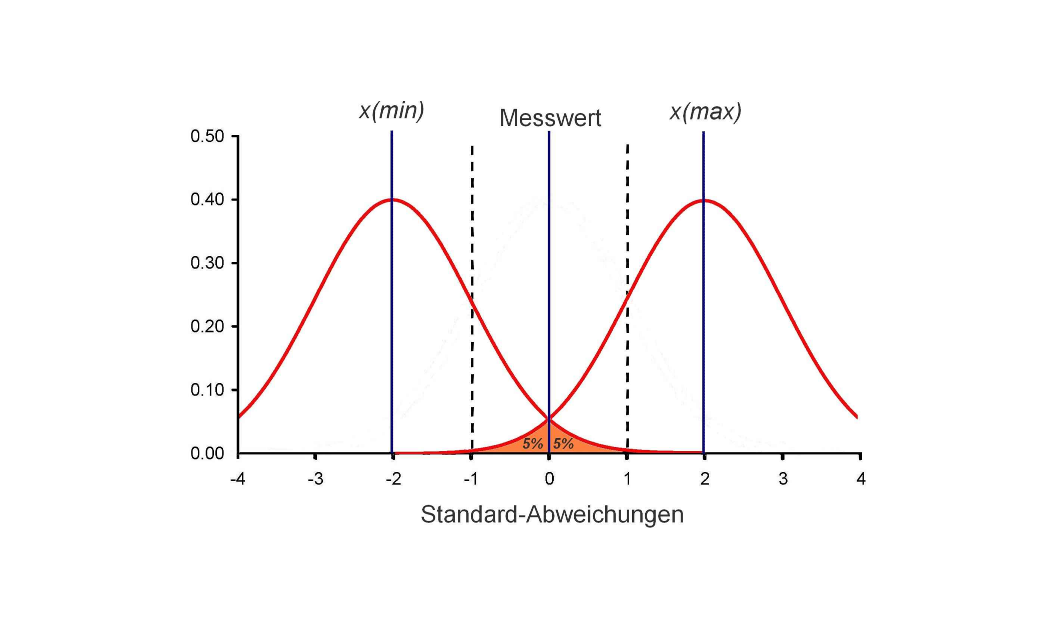


Heraklit

diskret	**kontinuierlich**
$P(n,k)$	$p(x)=\frac{1}{n}\cdot\frac{dn}{dx}$
$\sum_{i=1}^{n} p_i=1$	$\int p(x)\,dx=1$
$\mu=\sum_{i=1}^{n} p_i\cdot x_i$	$\mu=\int p(x)\cdot x\,dx$
$V=\sum_{i=1}^{n} p_i\cdot(x_i-\mu)^2$	$V=\int p(x)\cdot(x-\mu)^2\,dx$

Kontinuierliche Verteilungen

Bestimmung von Vertrauensintervallen



$$\int_{x_i=-\infty}^{x_{exp.}} p(x, x_{max})\, dx < y/2$$

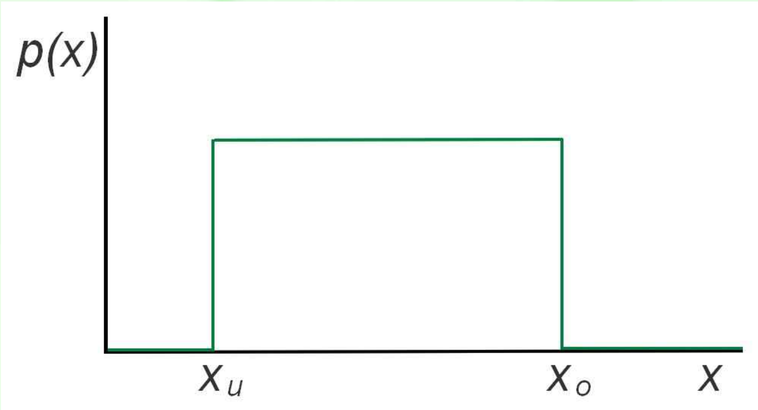
$$\int_{x_i=-\infty}^{x_{exp.}} p(x, x_{min})\, dx > 1 - y/2$$

Der Messwert liegt mit 95% Wahrscheinlichkeit oberhalb x(min)
Mit 90% Wahrscheinlichkeit zwischen x(min) und x(max)

Kontinuierliche Verteilungen

Münster University of Applied Sciences

Die Gleichverteilung



$$p(x) = konst. = \frac{1}{x_o - x_u}$$

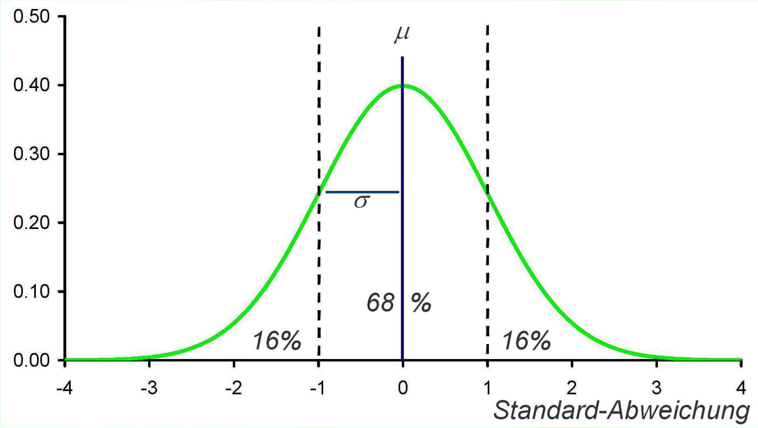
$$\mu = (x_o - x_u)/2$$

$$V = (x_o - x_u)^2/12$$

Beispiele: Zufallszahl zwischen 0 & 1
Spannungen zwischen den LSB Grenzen (ADC)
...

Kontinuierliche Verteilungen

Die Gauß – oder Normalverteilung



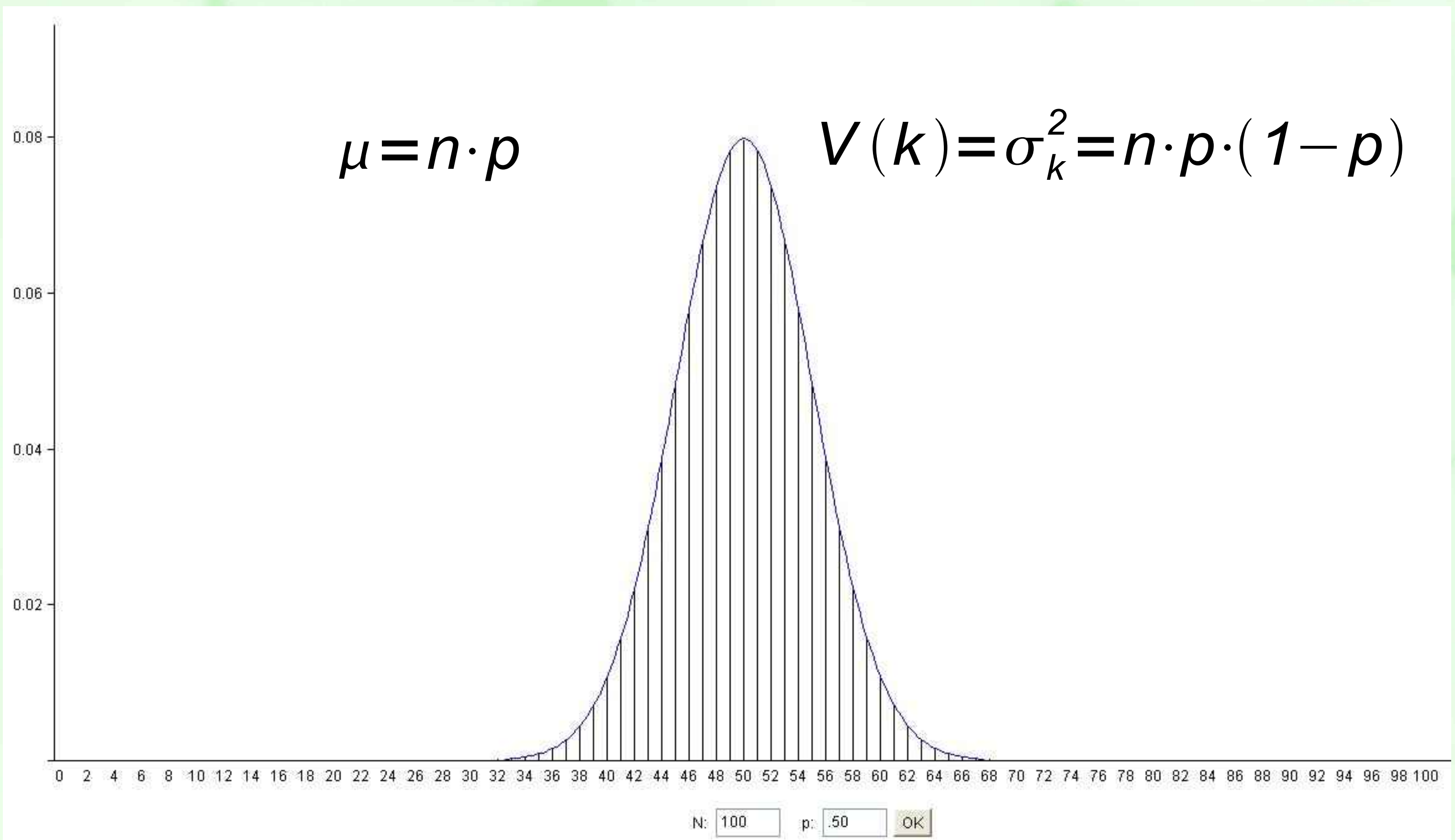
$$p(x) = \frac{1}{\sqrt{2\pi\sigma_x^2}} e^{-(x-\mu)^2/(2\sigma_x^2)}$$

„Central Limit Theorem“:
„alles wird Gauß“
(← viele Variable,
$p(x_i)$ egal)

Münster University of Applied Sciences

Kontinuierliche Verteilungen

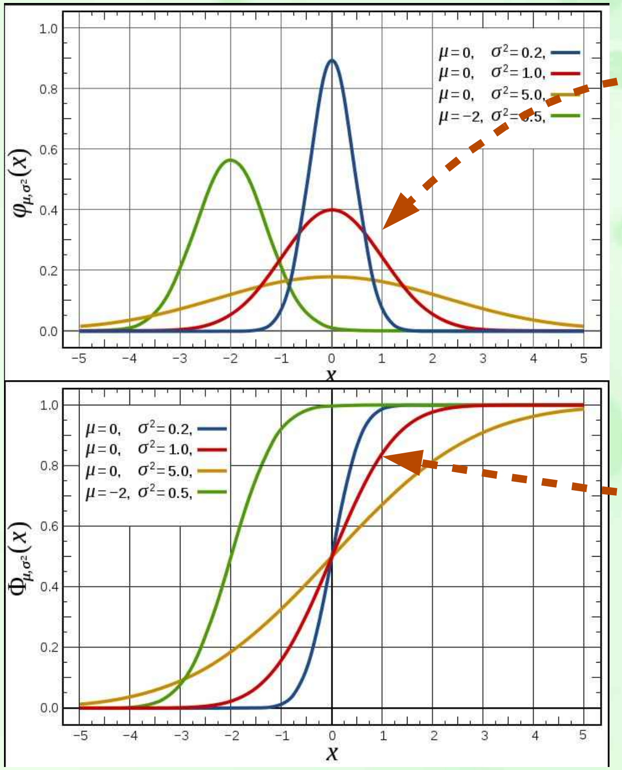
Gauß und Bernulli:
„man trifft sich“



Münster University of Applied Sciences

Kontinuierliche Verteilungen

Gauß: Bestimmung von Vertrauensintervallen



(Grafik aus Wikipedia)

$$p_{standard}(z) = \frac{1}{\sqrt{2\pi}} e^{-z^2/2}$$

$$z := (x-\mu)/\sigma \rightarrow dz = \frac{1}{\sigma} dx$$

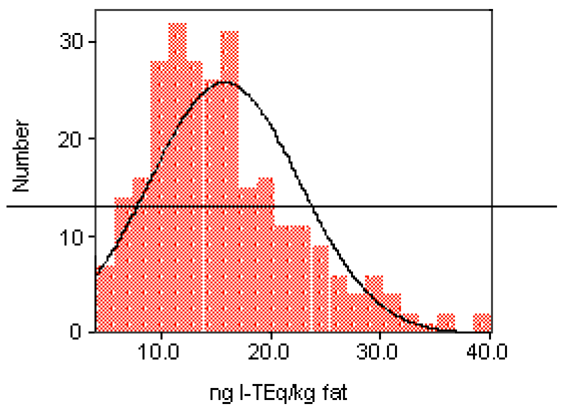
$$\Phi(z) = \int_{-\infty}^{z} p_{standard}(t)\, dt = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{z} e^{-t^2/2} dt$$

$$\int_{-\infty}^{x} p(t)\, dt = \Phi\left(\frac{x-\mu}{\sigma}\right)$$

Münster University of Applied Sciences

Kontinuierliche Verteilungen

Gauß: Bestimmung von Parametern (1)

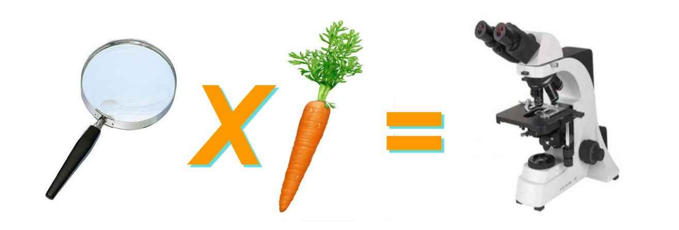


(aus einer Studie der WHO)

$$\sigma_x \approx \sigma_{(exp.)} = \sqrt{\frac{(x-\overline{x})^2}{n-1}}$$

$$\sigma_\mu = \frac{\sigma_x}{\sqrt{n}} \approx \sqrt{\frac{(x-\overline{x})^2}{n\cdot(n-1)}}$$

Für σ_μ gilt das Wurzelgesetz!

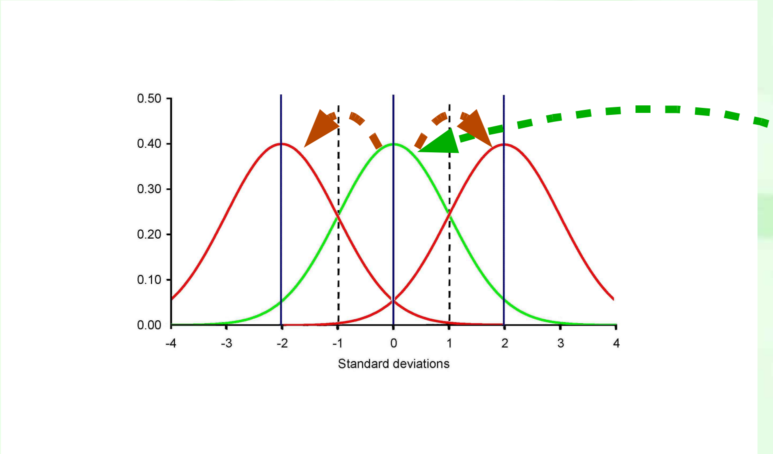


...aber nur, wenn es passt...

Münster University of Applied Sciences

Kontinuierliche Verteilungen

Gauß: Bestimmung von Parametern (2)



$$\sigma_x \approx \sigma_{(exp.)} = \sqrt{\frac{(x-\bar{x})^2}{n-1}}$$

$$\sigma_\mu = \frac{\sigma_x}{\sqrt{n}} \approx \sqrt{\frac{(x-\bar{x})^2}{n\cdot(n-1)}}$$

Voraussetzungen: Verteilungsfunktion symmetrisch
Verteilungsfunktionen für Werte neben μ genau so breit wie um μ herum

eigentlich eine Hypothesentest – aber man merkt es nicht...

Mehr Dimensionen!

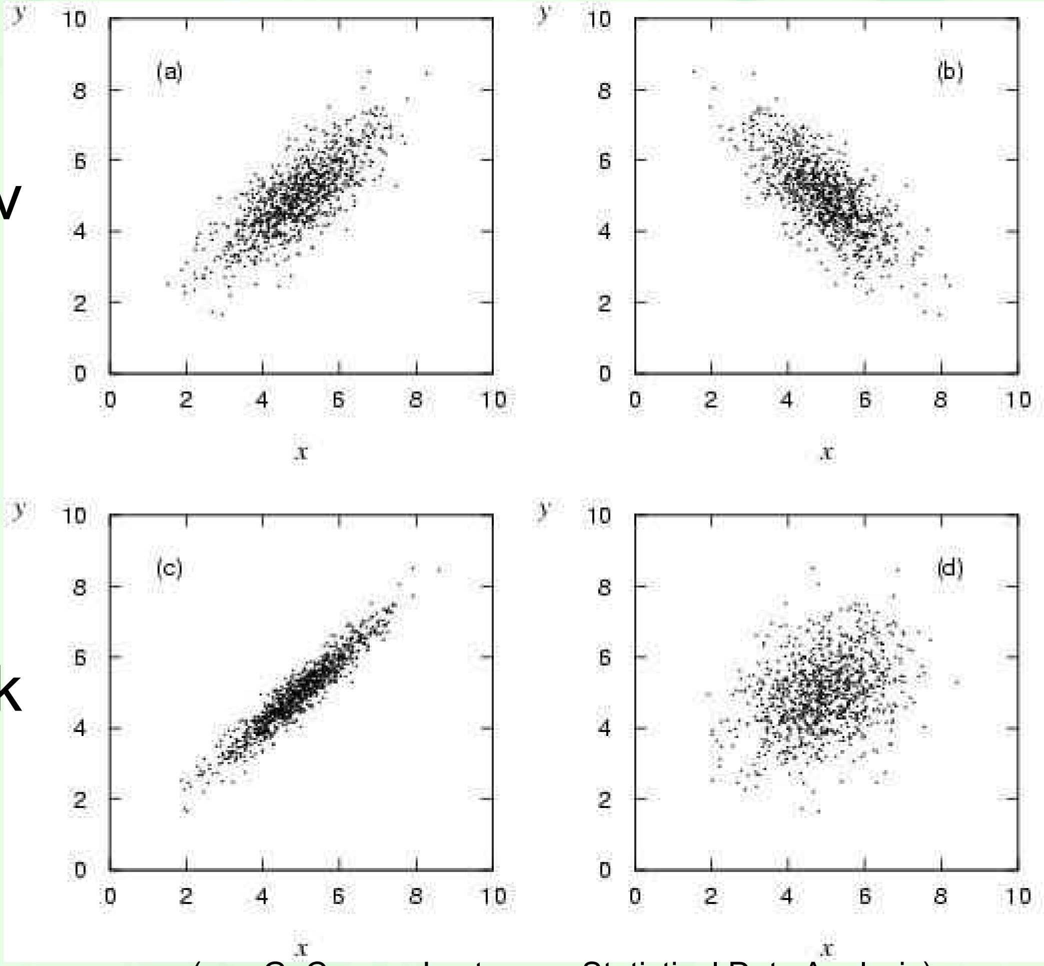
Das Grundproblem: Korrelationen

positiv

negativ

stark

schwach



(aus G. Cowan, Lecturs on Statistical Data Analysis).

Mehr Dimensionen!

Erweiterung der Begriffe

Varianz → **Kovarianz**

$$V(x)=E\left((x-\mu)^2\right) \qquad cov(x,y)=E\left((x-\mu_x)\cdot(y-\mu_y)\right)$$

Korrelationskoeffizient $\rho_{xy}=\dfrac{cov(x,y)}{\sigma_x\cdot\sigma_y}$

$$x_i=a\cdot y_i\rightarrow\rho_{xy}=1$$

$$x_i=-a\cdot y_i\rightarrow\rho_{xy}=-1$$

$$x_i\ unabhängig\ von\ y_i\rightarrow\rho_{xy}=0$$

Gilt der Umkehrschluss auch?

Mehr Dimensionen!

Fehlerfortpflanzung(1)

Wie kommt man vom Fehler der Messgrößen zur Fehler der Zielgröße?

Gegeben: $y(x_i) \qquad V_{i,j}(x_i, x_j) = cov(x_i, x_j)$

Gesucht: $V(y(x_j))$

...analytisch praktisch nicht zu schaffen!

→ **Potenzreihen um die Erwartungswerte:**

$$y(x_i) \approx y(\mu_i) + \sum_{i=1}^{n} \left[\frac{\partial y}{\partial x_i}\right]_{x_i=\mu_i} \cdot (x_i - \mu_i)$$

Mehr Dimensionen!

Fehlerfortpflanzung(2)

→ Potenzreihen um die Erwartungswerte:

$$y(x_i) \approx y(\mu_i) + \sum_{i=1}^{n} \left[\frac{\partial y}{\partial x_i}\right]_{x_i=\mu_i} \cdot (x_i - \mu_i)$$

und Verschiebungsgesetz

$$V(y) = E(y^2) - (E(y))^2 \approx E(y^2) - (E(\mu))^2$$

Potenzreihen in das Verschiebungsgesetz → …

→

$$\sigma_y^2 = \sum_{i,k=1}^{n} \left[\frac{\partial y \, \partial y}{\partial x_i \, \partial x_k}\right]_{x_i=\mu_i} \cdot V_{i,k}$$

Fehlerfortpflanzungsgesetz in *n* Dimensionen

Mehr Dimensionen!

Fehlerfortpflanzung(3)

Vektor-wertige Funktionen

Gegeben: $y_r(x_i) \quad V_{i,j}(x_i, x_j) = cov(x_i, x_j)$

Gesucht: $V_{r,s}(y(x_j))$

$$\rightarrow \qquad V_{r,s}(y_r, y_s) = \sum_{i,k=1}^{n} \left[\frac{\partial y_r \partial y_s}{\partial x_i \partial x_k} \right]_{x_i=\mu_i} \cdot V_{i,k}$$

oder, eleganter:

$$V(\vec{y}) = A \cdot V(\vec{x}) \cdot A^T, \quad A_{i,k} = \left[\frac{\partial y_i}{\partial x_k} \right]_{x_i=\mu_i}$$

Mehr Dimensionen!

Spezialfälle (1)

Differenz zweier Messwerte

$$y = x_1 - x_2 \rightarrow \sigma_y^2 = \sigma_{X1}^2 - 2\,cov(x_{1,}\, x_2) + \sigma_{X2}^2$$

→ unkorreliert $\rho_{x1\,x2} = 0 \quad \rightarrow \quad \sigma_y^2 = 2 \cdot \sigma_X^2$

→ antikorreliert $\rho_{x1\,x2} = -1 \quad \rightarrow \quad \sigma_y^2 = 4 \cdot \sigma_X^2$

→ 100% korreliert $\rho_{x1\,x2} = 1 \quad \rightarrow \quad \sigma_y^2 = 0\,(!!!)$

Folgen der Korrelationen bisweilen dramatisch

Mehr Dimensionen!

Spezialfälle (2)

Produkt zweier Messwerte

$$y = x_1 \cdot x_2 \rightarrow \left(\frac{\sigma_y}{\mu_y}\right)^2 = \left(\frac{\sigma_{X1}}{\mu_1}\right)^2 + 2\left(\frac{cov(x_{1,} x_2)}{\mu_1 \mu_2}\right) + \left(\frac{\sigma_{X2}}{\mu_2}\right)^2$$

→ unkorreliert für *n* Multiplikatoren:

$$y = x_1 \cdot x_2 \cdot \ldots x_n \rightarrow \left(\frac{\sigma_y}{\mu_y}\right)^2 = \sum_{i=1}^{n} \left(\frac{\sigma_{Xi}}{\mu_i}\right)^2$$

„Die Summe der quadrierten relativen Fehler“

Mehr Dimensionen!

Zusammenfassung

Korrelationen im Allgemeinen wichtig
→ Immer die gesamte Kovarianzmatrix benutzen

Brauche lineare Näherungen
→ funktioniert nur bei kleinen relativen Fehlern

Verfahren beinhaltet viele Parameter
(n Ableitungen, n Mittelwerte, n^2 Matrixelemente)
→ deutlich mehr als $n^2 + 2n$ Messwerte benutzen!

Mehr Freizeit!

Jetzt wäre ein günstiger Zeitpunkt zum Aufwachen