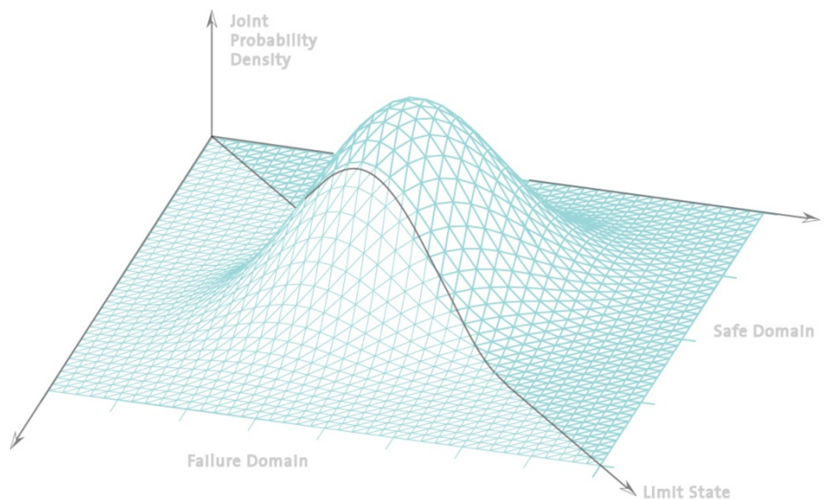


Statistik und Wahrscheinlichkeitsrechnung

Dr. Jochen Köhler

Änderung Übungsstunde

Die Gruppe von Markus trifft sich am Donnerstag statt im HCI D2 zusammen mit der Gruppe von Eva im HIL E4.

Kurze Zusammenfassung der letzten Vorlesung

Konfidenzintervall:

Typische Konfidenzintervalle und die dazugehörigen Fraktilwerte der Standardnormalverteilung:

Konvidenz	α	$k_{\alpha/2}$	$k_{1-\alpha/2}$
99%	0.01	-2.576	2.576
95%	0.05	-1.960	1.960
90%	0.1	-1.645	1.645

Inhalt der heutigen Vorlesung

- Kurze Zusammenfassung der letzten Vorlesung
- Schätzung und Modellentwicklung:
  - Methode der Momente
  - Methode der Maximum Likelihood
- Befragung der Studierenden (ca. 15 min)

Kleine Denkaufgabe 11.1

- Das Signifikanzniveau eines Hypothesentests entpricht:

der Wahrscheinlichkeit eines Fehlers 1. Art

der Wahrscheinlichkeit eines Fehlers 2. Art

je nach Wahl der Null-Hypothese entweder der Wahrscheinlichkeit eines Fehlers 1. Art oder eines Fehlers 2. Art.

Kleine Denkaufgabe 11.1 – Lösung

Das Signifikanzniveau eines Hypothesentests entpricht per Definition…

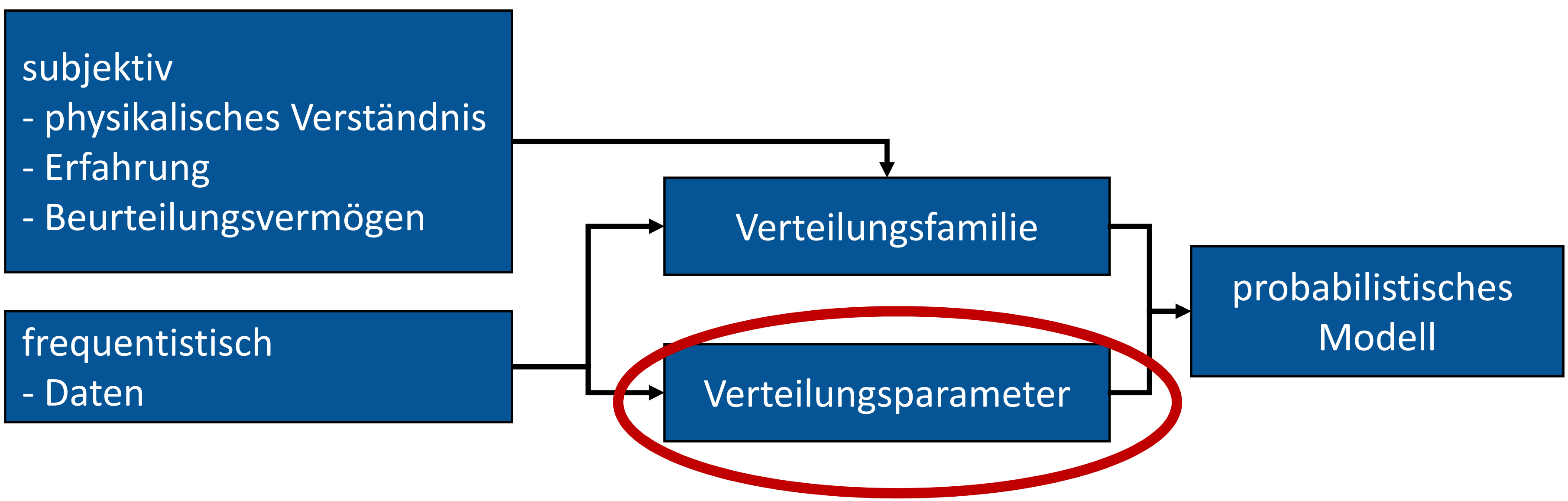
… der Wahrscheinlichkeit eines Fehlers 1. Art

Wahrheit / Urteil	H_0 trifft zu.	H_0 trifft nicht zu.
Akzeptanz von H_0.	Richtiges Urteil.	Fehler 2. Art.
Ausschluss von H_0.	Fehler 1. Art.	Richtiges Urteil.

Schätzung und Modellentwicklung - Übersicht

Wenn man Modelle im Ingenieurbereich entwickeln möchte, müssen unterschiedliche Typen von Informationen herangezogen werden.

- subjektive Informationen
- frequentistische Informationen



Schätzung der Verteilungsparameter

Haben wir uns für ein Verteilungstyp entschieden, müssen die Parameter abgeschätzt werden.

z.B. Normalverteilung

$$f_X(x) = \frac{1}{\sigma\sqrt{2\pi}} \exp\left(-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2 \right)$$

Weibullverteilung

$$f_X(x) = \frac{k}{u-\varepsilon}\left(\frac{x-\varepsilon}{u-\varepsilon}\right)^{k-1} \exp\left(-\left(\frac{x-\varepsilon}{u-\varepsilon}\right)^k \right)$$

Wahrscheinlichkeitsdichtefunktionen werden definiert durch ihre Parameter $\boldsymbol{\theta} = (\theta_1, \theta_2, .., \theta_k)^T$.

Allgemein gibt man Dichtefunktionen bedingt auf die Parameter an: $f_X(x|\boldsymbol{\theta})$

Schätzung der Verteilungsparameter

Es gibt eine Vielzahl von Methoden Verteilungsparameter zu schätzen; generell wird unterschieden zwischen:

- Punktschätzern
- Intervallschätzern.

Im folgenden werden wir zwei Methoden näher betrachten:

- Methode der Momente
- Methode der Maximum Likelihood

Schätzung der Verteilungsparameter

- Methode der Momenten (MoM)

Gegeben ist eine Stichprobe anhand dessen wir die Verteilungsparameter abschätzen: $\hat{\mathbf{x}} = (\hat{x}_1, \hat{x}_2, .., \hat{x}_n)^T$

Das Prinzip der Methode der Momenten ist die Parameter so abzuschätzen, dass die Momente der Verteilung und die Momente der Stichprobe identisch sind.

$$m_j = \frac{1}{n}\sum_{i=1}^{n} x_i^j$$

Stichprobe

$$\lambda_j = \lambda_j(\theta_1, \theta_2, .., \theta_k) = \int_{-\infty}^{\infty} x^j \cdot f_X(x|\boldsymbol{\theta})dx$$

Verteilung

Schätzung der Verteilungsparameter

- Methode der Momenten (MoM)

Angenommen wir haben eine Verteilung mit k Parametern:

Dann haben wir k Gleichungen mit k unbekannten zu lösen:

$$m_j = \lambda_j(\boldsymbol{\theta}),\ j = 1, 2, .., k$$

$$\Downarrow$$

$$\boxed{\frac{1}{n}\sum_{i=1}^{n} x_i^j} = \boxed{\int_{-\infty}^{\infty} x^j \cdot f_X(x|\boldsymbol{\theta})dx}, \qquad j = 1, 2, .., k$$

Stichprobe

Verteilung

Schätzung der Verteilungsparameter

- Methode der Momenten (MoM)

Beispiel: Beton Druckfestigkeit, Annahme Normalverteilt:

Die Normalverteilung hat zwei Parameter – wir müssen also zwei Gleichungen lösen:

$$\boxed{m_1 = \frac{1}{n}\sum_{i=1}^{n}\hat{x}_i} = \boxed{\lambda_1 = \int_{-\infty}^{\infty} x \cdot f_X(x|\mu,\sigma)dx}$$

$$\boxed{m_2 = \frac{1}{n}\sum_{i=1}^{n}\hat{x}_i^2} = \boxed{\lambda_2 = \int_{-\infty}^{\infty} x^2 \cdot f_X(x|\mu,\sigma)dx}$$

Betondruck-festigkeit [Mpa]
24.4
27.6
27.8
27.9
28.5
30.1
30.3
31.7
32.2
32.8
33.3
33.5
34.1
34.6
35.8
35.9
36.8
37.1
39.2
39.7

Schätzung der Verteilungsparameter

- Methode der Momenten (MoM)

Die Stichprobenmomente sind:

$$m_1 = \frac{1}{20}\sum_{i=1}^{n} \hat{x}_i = 32.67 \qquad m_2 = \frac{1}{20}\sum_{i=1}^{n} \hat{x}_i^2 = 1083.36$$

Die Momente der Verteilung sind:

$$\lambda_1 = \int_{-\infty}^{\infty} x \cdot \frac{1}{\sigma\sqrt{2\pi}} \exp\left(-0.5\frac{(x-\mu)^2}{\sigma^2}\right)dx \qquad \lambda_2 = \int_{-\infty}^{\infty} x^2 \cdot \frac{1}{\sigma\sqrt{2\pi}} \exp\left(-0.5\frac{(x-\mu)^2}{\sigma^2}\right)dx$$

Schätzung der Verteilungsparameter

- Methode der Momenten (MoM)

Bei Formulierung der folgenden Objektfunktion:

$$g(\mu,\sigma)=(\lambda_1(\mu,\sigma)-m_1)^2+(\lambda_2(\mu,\sigma)-m_2)^2$$

Lassen sich die Parameter durch Optimierung finden.

z.B. mit Hilfe von MS EXCEL.

Schätzung der Verteilungsparameter

- Methode der Momenten (MoM)

ODER:

Alternativ und viel einfacher ist die Herangehensweise mithilfe von den **zentralen** Momenten.

1. Bestimmen des Stichprobenmittelwertes und der Stichprobenstandardabweichung.
2. Gleichsetzen mit dem Stichprobenmittelwert und der Standardabweichung der Verteilungsfunktion.
3. Errechnen der Parameter.

Schätzung der Verteilungsparameter

- Methode der Momenten (MoM)

Log-Normalverteilung,	Parameter	Momente
$F_X(x) = \Phi\left(\frac{\ln(x)-\lambda}{\zeta}\right)$ $f_X(x) = \frac{1}{x\zeta\sqrt{2\pi}}\exp\left(-\frac{1}{2}\left(\frac{\ln(x)-\lambda}{\zeta}\right)^2\right)$	$0 \leq x \leq \infty$ λ, ζ $\zeta > 0$	$\mu = \exp\left(\lambda + \frac{\zeta^2}{2}\right)$ $\sigma = \mu\sqrt{\exp(\zeta^2)-1}$

Gumbel max.	Parameter	Momente
$f_X(x) = \alpha\exp(-\alpha(x-u)-\exp(x-u))$ $F_X(x) = \exp(-\exp(-\alpha(x-u)))$	$-\infty \leq x \leq \infty$ u $\alpha > 0$	$\mu = u + \frac{\gamma}{\alpha}$, $\gamma \approx 0.5772$ $\sigma = \frac{\pi}{\alpha\sqrt{6}}$

Kleine Denkaufgabe 11.2

- Welche der folgenden Aussagen in Bezug auf Hypothesentests ist richtig?

Die Wahl der Null Hypothese ist subjektiv.

Das Signifikanzniveau sollte so klein wie möglich ausgewählt werden.

Wenn die Anzahl der Stichproben bekannt ist, kann das Signifikanzniveau beliebig gewählt werden.

Kleine Denkaufgabe 11.2 – Lösung

Die Wahl der Null Hypothese ist subjektiv.

Auf gleiche Weise kann auch die Wahl des Signifikanzniveaus subjektiv sein.

Das subjektive Auswählen von Null Hypothesen und Signifikanzniveaus kann als **Prozess der Entscheidungsfindung** verstanden werden, der die ökonomischen Auswirkungen von **Konsequenzen** bedingt durch statistische Schlussfolgerungen mit einbezieht.

Schätzung der Verteilungsparameter

- Maximum Likelihood Methode (MLM)

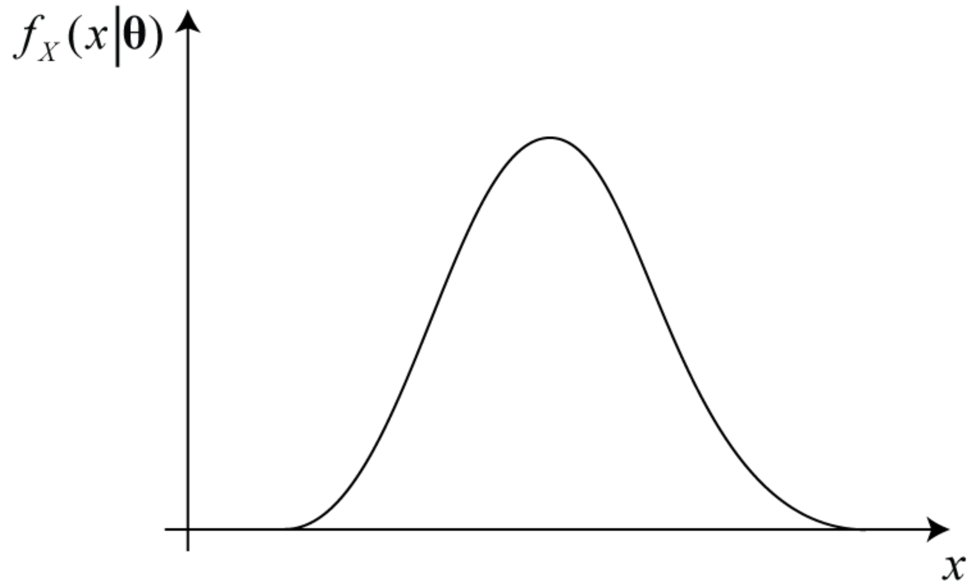
Die Grundidee der MLM ist

die Parameter der Verteilung zu identifizieren, welche die maximale ‚Likelihood' besitzen – welche also am wahrscheinlichsten sind.

Schätzung der Verteilungsparameter

- Maximum Likelihood Methode (MLM)

Die Form einer Wahrscheinlichkeitsdichtefunktion ist bestimmt durch ihre Parameter $\boldsymbol{\theta} = (\theta_1, \theta_2, .., \theta_k)^T$.



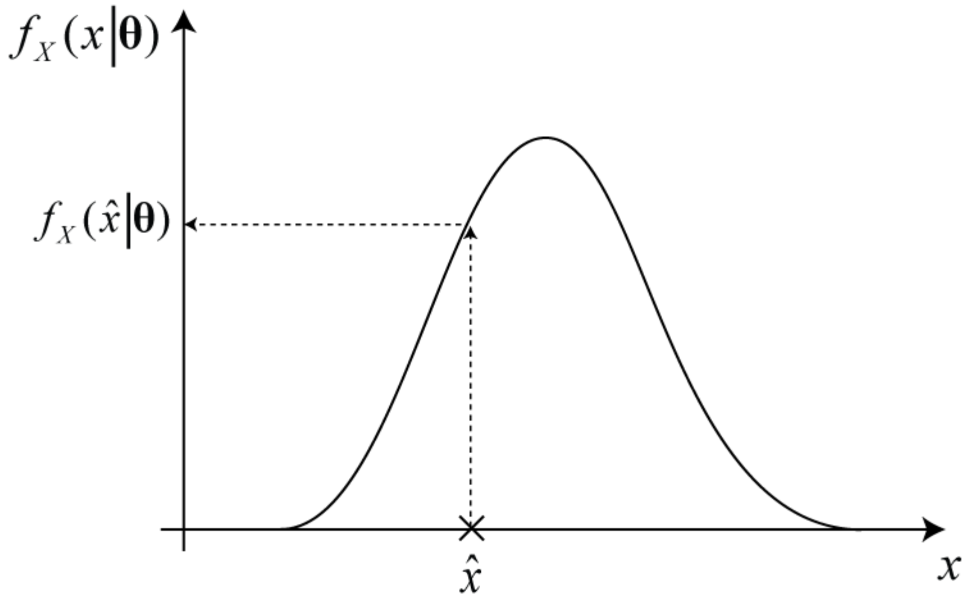
Schätzung der Verteilungsparameter

- Maximum Likelihood Methode (MLM)

Die Form einer Wahrscheinlichkeitsdichtefunktion ist bestimmt durch ihre Parameter $\boldsymbol{\theta} = (\theta_1, \theta_2, .., \theta_k)^T$.

Wird nun eine Stichprobe $\hat{\mathbf{x}} = (\hat{x}_1, \hat{x}_2, .., \hat{x}_n)^T$ beobachtet, hat eine Realisation der Stichprobe $\hat{x}$ die ‚Likelihood':

$$L = f_X(\hat{x} | \boldsymbol{\theta})$$



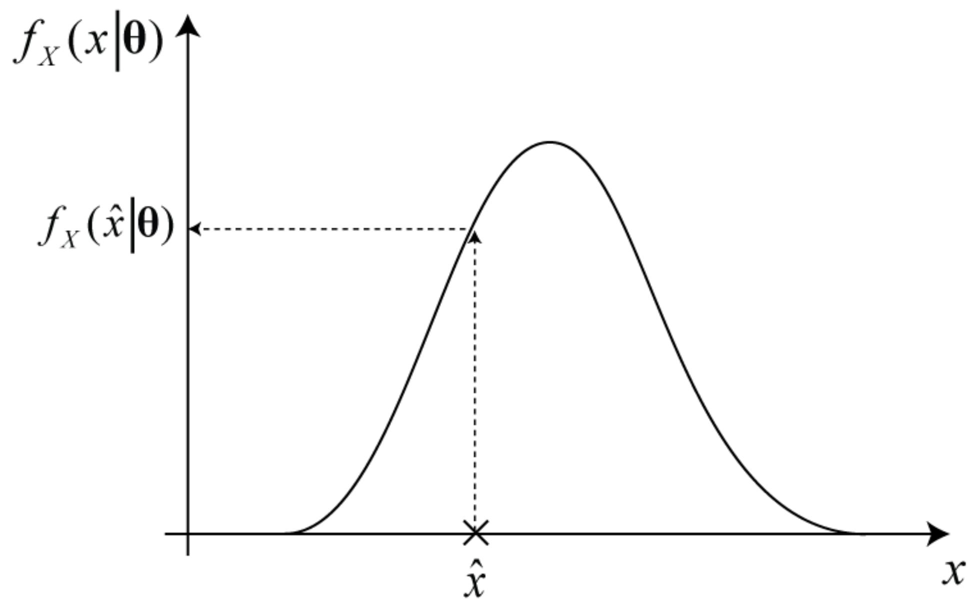
Schätzung der Verteilungsparameter

- Maximum Likelihood Methode (MLM)

Die Form einer Wahrscheinlichkeitsdichtefunktion ist bestimmt durch ihre Parameter $\boldsymbol{\theta} = (\theta_1, \theta_2, .., \theta_k)^T$.

Wird nun eine Stichprobe $\hat{\mathbf{x}} = (\hat{x}_1, \hat{x}_2, .., \hat{x}_n)^T$ beobachtet, hat eine Realisation der Stichprobe $\hat{x}$ die ‚Likelihood':

$$L = f_X(\hat{x} | \boldsymbol{\theta})$$



Die ‚Likelihood' der gesamten Stichproben errechnet sich aus dem Produkt der Likelihoods der Einzelbeobachtungen -> $L = \prod_{i=1}^{n} f_X(\hat{x}_i | \boldsymbol{\theta})$

-> bei gegebenen Parametern

Schätzung der Verteilungsparameter

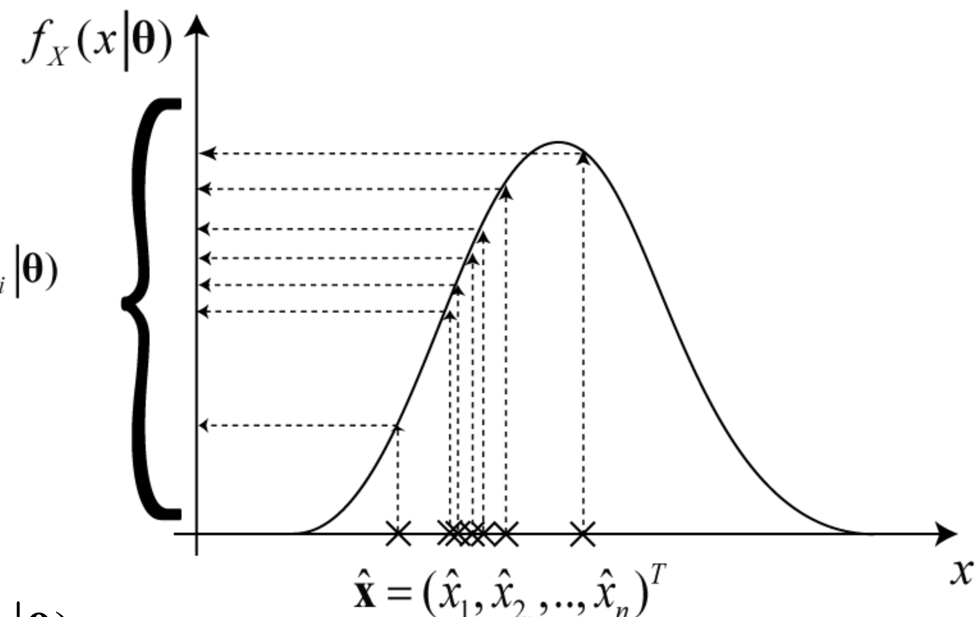
- Maximum Likelihood Methode (MLM)

Die Form einer Wahrscheinlichkeitsdichtefunktion ist bestimmt durch ihre Parameter $\boldsymbol{\theta} = (\theta_1, \theta_2, .., \theta_k)^T$.

Wird nun eine Stichprobe $\hat{\mathbf{x}} = (\hat{x}_1, \hat{x}_2, .., \hat{x}_n)^T$ beobachtet, hat eine Realisation der Stichprobe $\hat{x}$ die ‚Likelihood‘:

$$L = f_X(\hat{x} | \boldsymbol{\theta})$$

$$L = \prod_{i=1}^{n} f_X(\hat{x}_i | \boldsymbol{\theta})$$



Die ‚Likelihood‘ der gesamten Stichproben errechnet sich aus dem Produkt der Likelihoods der Einzelbeobachtungen -> $L = \prod_{i=1}^{n} f_X(\hat{x}_i | \boldsymbol{\theta})$

-> bei gegebenen Parametern

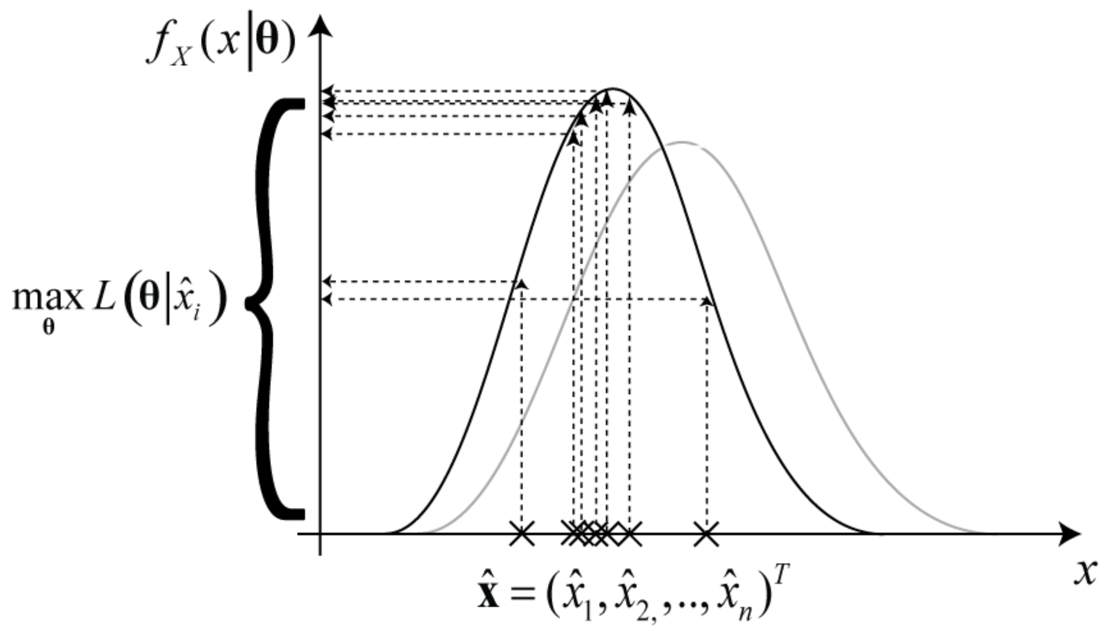
Schätzung der Verteilungsparameter

- Maximum Likelihood Methode (MLM)

Wir maximieren die ‚Likelihood' unter Veränderung der Parameter:

$$\max_{\boldsymbol{\theta}} L\left(\boldsymbol{\theta} \middle| \hat{x}_i\right)$$

-> und erhalten die Parameter welche am besten zu der Stichprobe passen.



Schätzung der Verteilungsparameter

- Maximum Likelihood Methode (MLM)

Beispiel Normalverteilung:
Wir nehmen an, die unseren Beobachtungen zugrundeliegende Verteilungsfunktion ist die Normalverteilung.

$$f_X(x) = \frac{1}{\sigma\sqrt{2\pi}} exp\left(-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2\right)$$

Die Likelihood der Beobachtungen der Stichprobe $\hat{\mathbf{x}} = (\hat{x}_1, \hat{x}_2, \ldots, \hat{x}_n)^T$ ist dann;

$$L(\boldsymbol{\theta}|\hat{\mathbf{x}}) = \prod_{i=1}^{n} \frac{1}{\sigma\sqrt{2\pi}} \exp\left(-\frac{1}{2}\left(\frac{\hat{x}_i-\mu}{\sigma}\right)^2\right)$$

Schätzung der Verteilungsparameter

- Maximum Likelihood Methode (MLM)

Beispiel Normalverteilung:
Die Parameter der Normalverteilung $\boldsymbol{\theta}$ werden so gewählt, dass sie die Likelihoodfunktion maximieren:

$$\min_{\boldsymbol{\theta}} \left(-L\left(\boldsymbol{\theta} \middle| \hat{\mathbf{x}}\right)\right)$$

-> es ist von Vorteil die logarithmierte Likelihoodfunktion zu verwenden:

$$l(\boldsymbol{\theta}|\mathbf{x}) = \sum_{i=1}^{n} \ln(f_X(\hat{x}_i|\boldsymbol{\theta})) \qquad -l(\boldsymbol{\theta}|\mathbf{x}) = -\sum_{i=1}^{n} \ln(f_X(\hat{x}_i|\boldsymbol{\theta}))$$

Schätzung der Verteilungsparameter

- Maximum Likelihood Methode (MLM)

Beispiel Normalverteilung:
Werden die Parameter $\boldsymbol{\theta}$ so gewählt, dass sie die log-Likelihoodfunktion maximieren:

$$\min_{\boldsymbol{\theta}}\left(-l(\boldsymbol{\theta}\,|\,\hat{\mathbf{x}})\right)$$

Kann gezeigt werden, dass die Parameter an sich zu Normalverteilten Zufallsvariablen konvergieren:

Mit Mittelwerten: $\boldsymbol{\mu}_{\Theta} = (\theta_1^*, \theta_2^*, .., \theta_n^*)^T$

Und Kovarianzmatrix: $\mathbf{C}_{\Theta\Theta} = \mathbf{H}^{-1}$ wobei $H_{ij} = \left.\dfrac{\partial^2 - l(\boldsymbol{\theta}|\hat{\mathbf{x}})}{\partial\theta_i \partial\theta_j}\right|_{\boldsymbol{\theta}=\boldsymbol{\theta}^*}$

Schätzung der Verteilungsparameter

- Maximum Likelihood Methode (MLM)

Beispiel Normalverteilung:
Betrachten wir nun die Daten der Betondruckfestigkeit, unter der Annahme, dass die Betondruckfestigkeit einer Normalverteilung folgt, ergibt sich für die log-Likelihoodfunktion:

$$l(\boldsymbol{\theta}|\hat{\mathbf{x}}) = n \cdot \ln\left(\frac{1}{\sqrt{2\pi}\theta_1}\right) - \frac{1}{2}\sum_{i=1}^{n}\frac{(\hat{x}_i - \theta_2)^2}{\theta_1^2}$$

Das Minimum kann analytisch wie folgt bestimmt werden:

$$\frac{\partial l}{\partial \theta_1} = -\frac{n}{\theta_1} + \frac{1}{\theta_1^3}\sum_{i=1}^{n}(\hat{x}_i - \theta_2)^2 = 0$$

$$\frac{\partial l}{\partial \theta_2} = \frac{1}{\theta_1^2}\sum_{i=1}^{n}(\hat{x}_i - \theta_2) = 0$$

$$\Rightarrow$$

$$\theta_1 = \sqrt{\frac{\sum_{i=1}^{n}(\hat{x}_i - \theta_2)^2}{n}}$$

$$\theta_2 = \frac{1}{n}\sum_{i=1}^{n}\hat{x}_i$$

Schätzung der Verteilungsparameter

- Maximum Likelihood Methode (MLM)

Beispiel Normalverteilung:
Gemäss Zahlenbeispiel ‚Betondruckfestigkeit' ergibt sich:

$$\theta_1 = \sqrt{\frac{\sum_{i=1}^{n}\left(\hat{x}_i - \theta_2\right)^2}{n}} = \sqrt{\frac{367.19}{20}} = 4.05$$

Mittelwert der Standardabweichung – Biased (!)

$$\theta_2 = \frac{1}{n}\sum_{i=1}^{n}\hat{x}_i = \frac{653.3}{20} = 32.67$$

Mittelwert des Mittelwertes

Schätzung der Verteilungsparameter

- Maximum Likelihood Methode (MLM)

Beispiel Normalverteilung:
Gemäss Zahlenbeispiel ‚Betondruckfestigkeit' ergibt sich:

Für die Kovarianzmatrix:

$$H = \begin{pmatrix} \frac{n}{\theta_1} - \frac{3\sum_{i=1}^{n}(x_i - \theta_2)^2}{\theta_1^4} & \frac{2\sum_{i=1}^{n}(x_i - \theta_2)}{\theta_1^3} \\ \frac{2\sum_{i=1}^{n}(x_i - \theta_2)}{\theta_1^3} & \frac{n}{\theta_1^2} \end{pmatrix}$$

$$C_{\Theta\Theta} = H^{-1} = \begin{pmatrix} \boxed{0.836} & 0 \\ 0 & \boxed{0.165} \end{pmatrix}$$

Varianz des Mittelwertes

Varianz der Standardabweichung

Schätzung der Verteilungsparameter

- Maximum Likelihood Methode (MLM)

Das Problem lässt sich ebenfalls numerisch lösen

– z.B. mit MS EXCEL.

Schätzung der Verteilungsparameter

- Maximum Likelihood Methode (MLM)

Illustration