

Moderne Methoden der Datenanalyse – Parameterschätzung –

Roger Wolf
28. Mai 2020

Inhalt der nächsten (drei) Vorlesungen

- Grundlagen zur Parameterschätzung und Einführung von Begriffen und Nomenklatur.
- Parameterschätzung mit Hilfe der Maximum Likelihood Methode.
- Parameterschätzung mit und Besonderheiten der χ^2 -Methode.
- Parameterisierungs- und Minimierungsmethoden (evtl. nur zum Lesen).

Bedeutung von χ_{obs}^2 für die LS-Schätzung

Für normalverteilte Zufallsgrößen folgt der quadratische Abstand der LS-Schätzwerte $\hat{y}_i(\{x_i\}, \{\hat{\theta}_j\})$ von den Messwerten $\{y_i\}$ für $j = 1, \dots, k$

$$\chi_{\text{obs}}^2 = \sum_{i \leq n} \left(\hat{y}_i(\{x_i\}, \hat{\theta}_j) - y_i \right)^2$$

einer $\chi^2(x, n - k)$ -Verteilung mit $n - k$ Freiheitsgraden.

(siehe [Lecture_05 Folien 17ff](#))

- Dabei entspricht n der Anzahl der Messungen und k der Anzahl der Parameter zur Anpassung.
- Da der Erwartungswert von $\chi^2(x, n - k)$ $n - k$ ist, ist bei zugrundeiegender Normalverteilung der Messwerte

$$\frac{\chi_{\text{obs}}^2}{n - k} \approx 1$$

zu erwarten.

Bedeutung von χ_{obs}^2 für die LS-Schätzung

- Unter der Annahme, dass Ihren Messwerten eine Normalverteilung zugrundeliegt können Sie die folgenden Schlüsse ziehen:

$$\frac{\chi_{\text{obs}}^2}{n - k} \ll 1 \quad \Rightarrow \quad \text{Fehler zu gross abgeschätzt, oder zu viele Parameter im Modell?}$$

$$\frac{\chi_{\text{obs}}^2}{n - k} \gg 1 \quad \Rightarrow \quad \text{Fehler zu klein abgeschätzt, oder Modell falsch.}$$

Verteilungsfreiheit

- Beachten Sie, dass für die LS-Schätzung nur die Werte ($\{y_i\}$) und Unsicherheiten ($\{\sigma_i\}$) der Messreihe angegeben werden. Im Gegensatz zur ML-Schätzung gibt es keine Aussage über die zugrundeliegende Wahrscheinlichkeitsdichte der Einzelmessungen!
- Man bezeichnet diese Eigenschaft als Verteilungsfreiheit der LS-Schätzung.
- Wir demonstrieren diese Eigenschaft anhand des Beispiels aus [Lecture_05](#) (Folie 25, „Anpassung einer Geraden an fünf Messpunkte“):

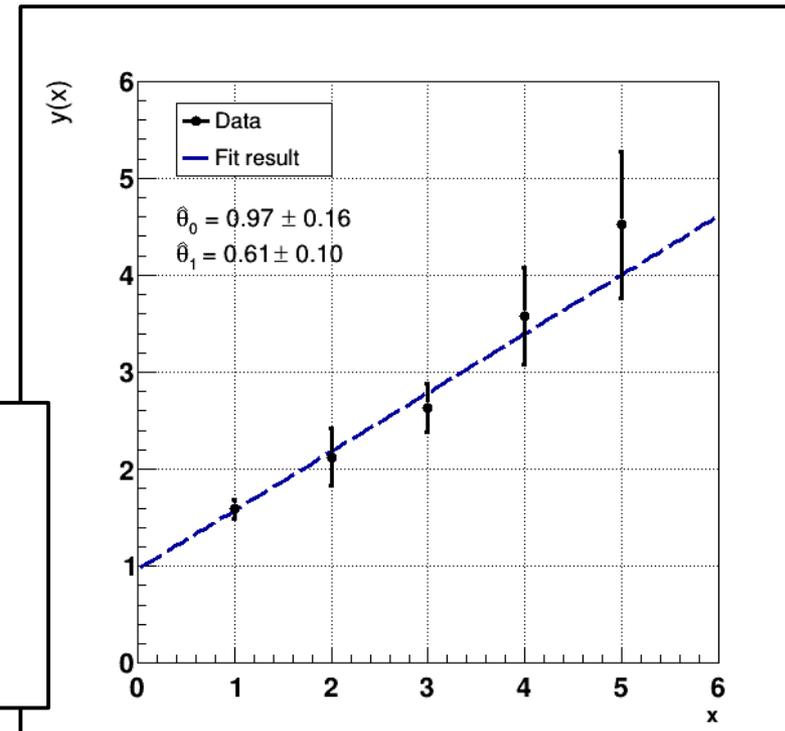
- Erinnerung:

Wahrheit:

$$\theta_0 = 1.0$$

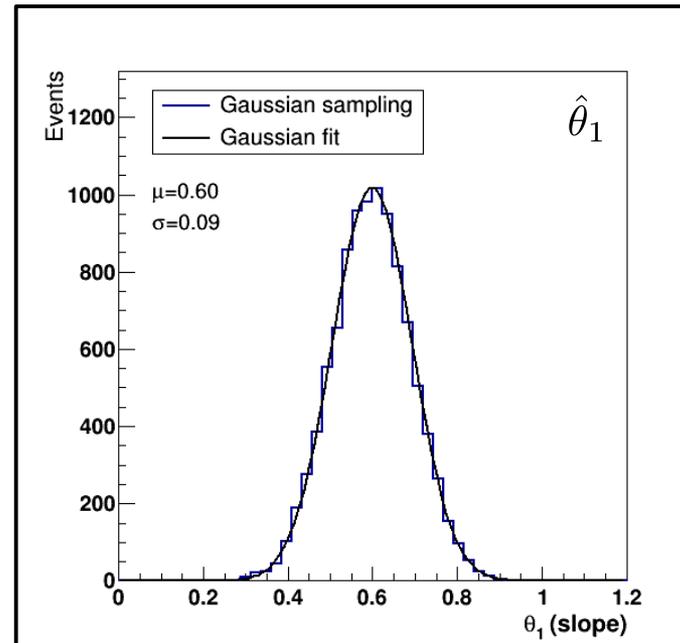
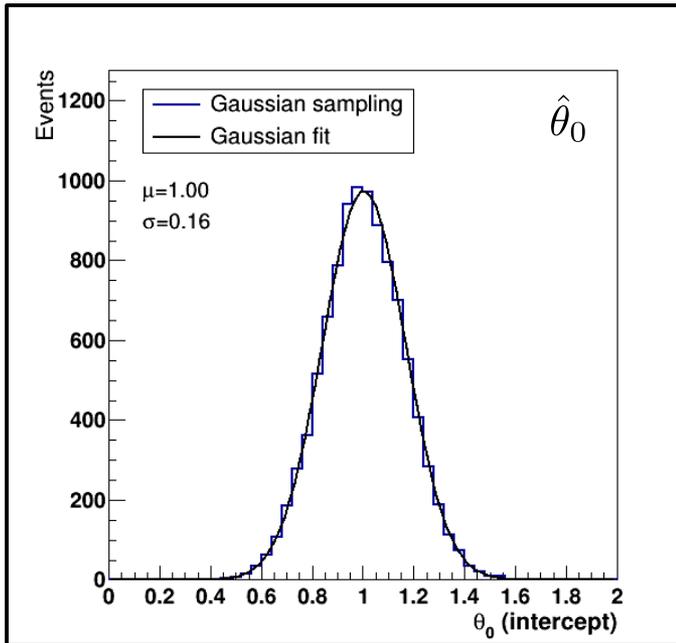
$$\theta_1 = 0.6$$

- Rechts sehen Sie einen möglichen Ausgang der Messung als Pseudoexperiment, wofür die Messwerte zufällig neuverteilt (randomisiert) wurden.
- Die zugrundeliegende Wahrscheinlichkeitsdichte für jede Einzelmessung war dabei vorgegeben durch $\varphi(x, y_i, \sigma_i)$.



Verteilungsfreiheit – zugrundeliegende Normalverteilung

- Hier sehen Sie die Ausgänge der Messung nach 10k-facher Wiederholung des Pseudoexperiments, wobei jede Einzelmessung nach $\varphi(x, y_i, \sigma_i)$ normalverteilt ist:

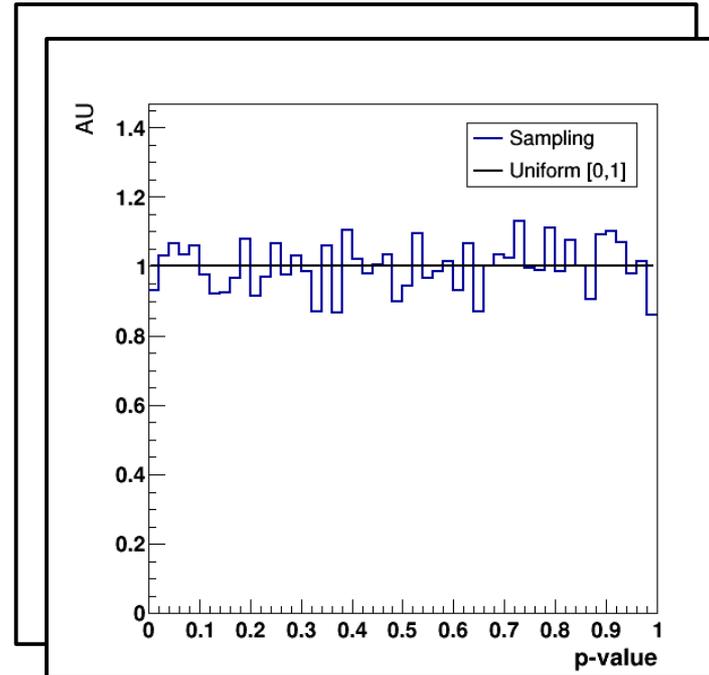
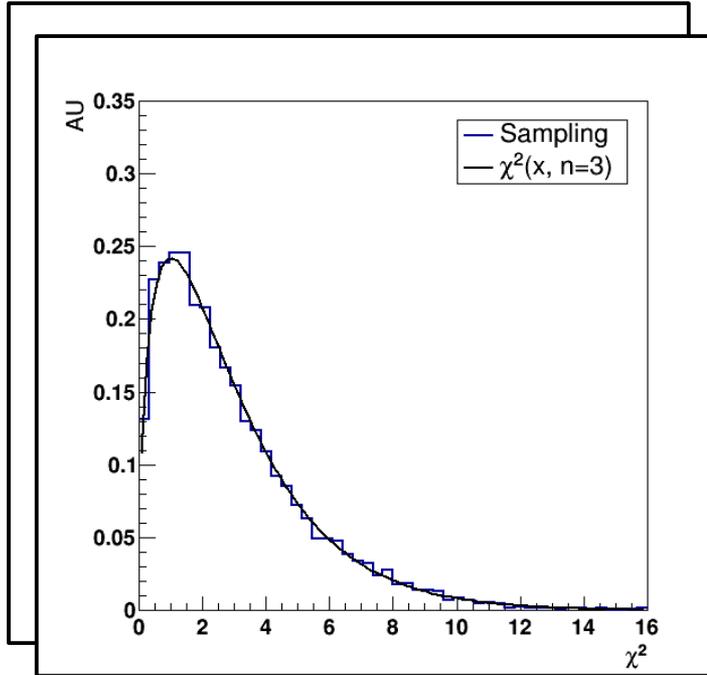


Ein lauffähiges ROOT macro finden Sie [hier](#).

- Die Verteilungen der LS-Schätzungen $\hat{\theta}_0$ und $\hat{\theta}_1$ sind jeweils normalverteilt und unverzerrt.
- NB:** Vergleichen Sie μ und σ mit [Lecture_05](#) Folie 27.

Verteilungsfreiheit – zugrundeliegende Normalverteilung

- Hier sehen Sie die Ausgänge der Messung nach 10k-facher Wiederholung des Pseudoexperiments, wobei jede Einzelmessung nach $\varphi(x, y_i, \sigma_i)$ normalverteilt ist:



Der quadratische Abstand der LS-Schätzwerte von den wahren Werten

$$\chi_{\text{obs}}^2 = \sum_{i \leq n} \left(\hat{y}_i(\{x_i\}, \hat{\theta}_j) - y_i \right)^2$$

ist nach $\chi^2(x, 3)$ verteilt.

Der p-Wert

$$p = \int_{\chi_{\text{obs}}^2}^{\infty} \chi^2(x, 3) dx$$

ist uniform verteilt.

Erinnerung p-Wert

- Wir erinnern hier nochmal durch indirekte Rechnung daran, dass der p-Wert einer zugrundeliegenden Verteilung gleichverteilt ist (hier am Beispiel der zugrundeliegenden $\chi^2(x, n)$ Verteilung):

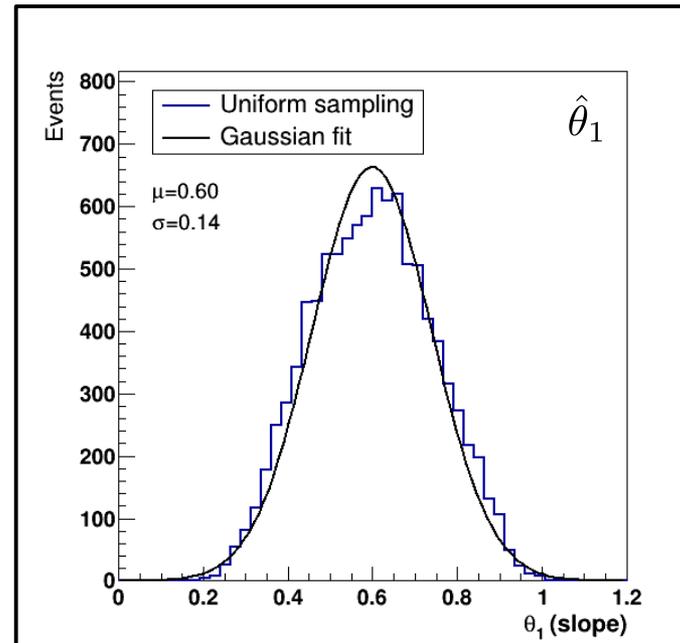
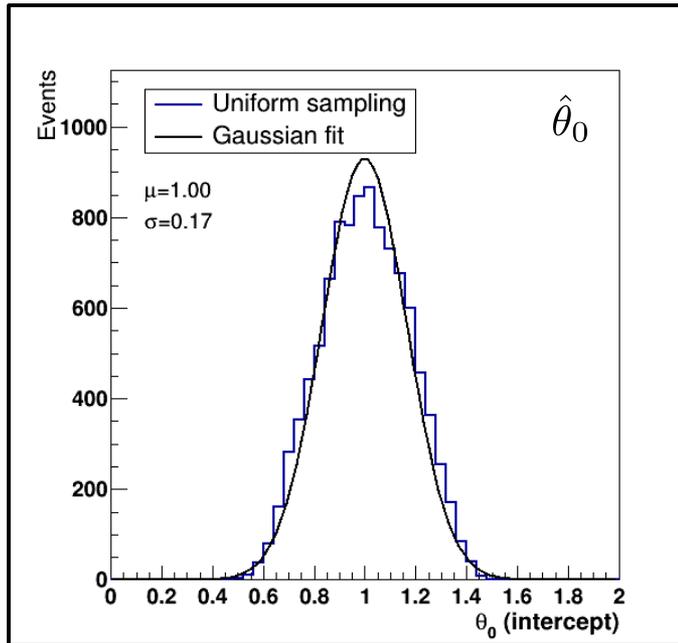
$$p(\chi^2 \geq \chi_{\text{obs}}^2) = \int_0^p 1 \, dp' = \int_{\chi_{\text{obs}}^2}^{\infty} \chi^2(x') \, dx'$$

p gleichverteilt.

χ_{obs}^2 wirklich nach
 $\chi^2(x, n)$ verteilt.

Verteilungsfreiheit – zugrundeliegende uniforme Verteilung

- Hier sehen Sie die Ausgänge der Messung nach 10k-facher Wiederholung des Pseudo-experiments, wobei jede Einzelmessung nach $U(x)$ *gleichverteilt* ist:

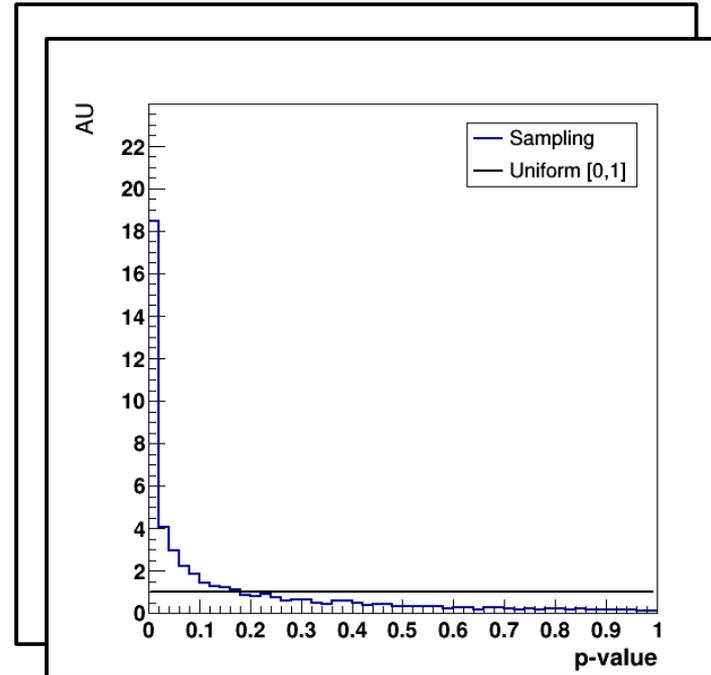
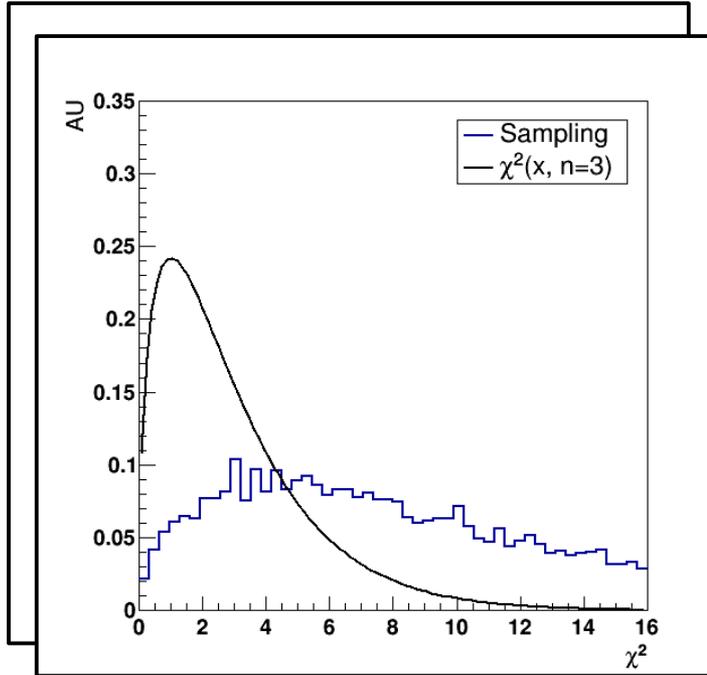


Ein lauffähiges ROOT macro finden Sie [hier](#).

- Die Verteilungen der LS-Schätzungen $\hat{\theta}_0$ und $\hat{\theta}_1$ sind jeweils normalverteilt und unverzerrt, aber nicht mehr effizient.

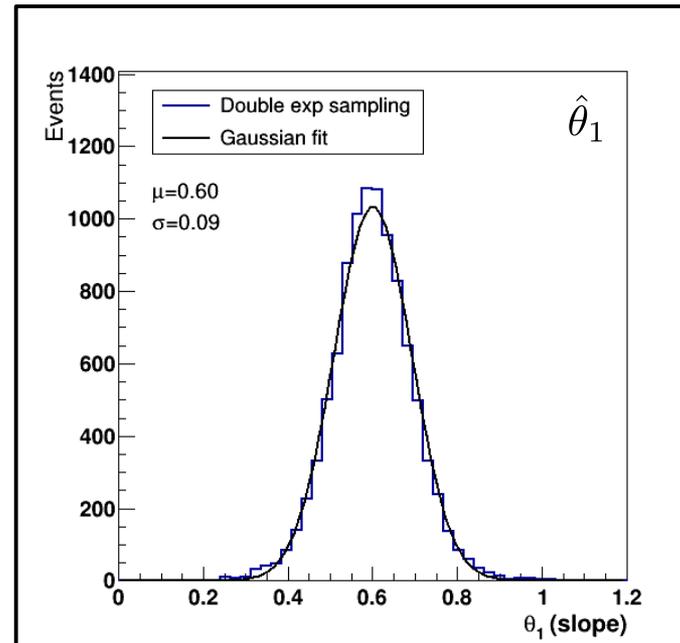
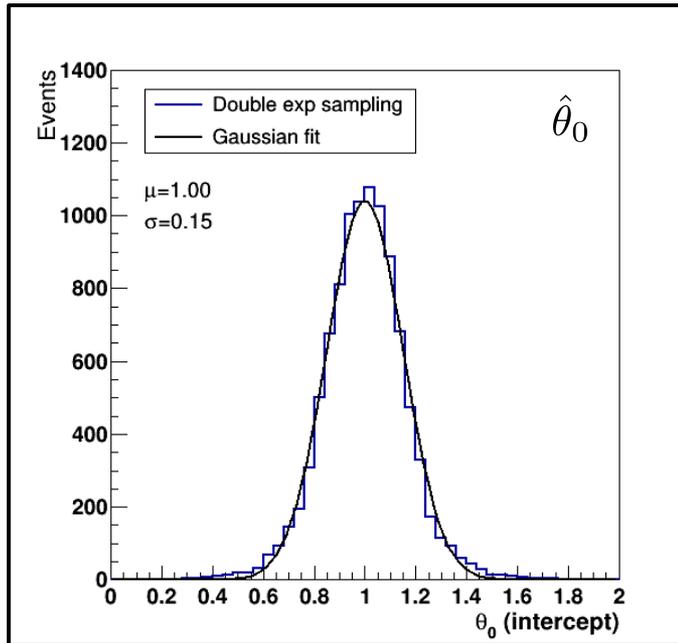
Verteilungsfreiheit – zugrundeliegende uniforme Verteilung

- Hier sehen Sie die Ausgänge der Messung nach 10k-facher Wiederholung des Pseudo-experiments, wobei jede Einzelmessung nach $U(x)$ gleichverteilt ist:



Verteilungsfreiheit – zugrundeliegende exp Verteilung

- Hier sehen Sie die Ausgänge der Messung nach 10k-facher Wiederholung des Pseudo-experiments, wobei jede Einzelmessung beidseitig nach $\exp(x, \sigma_i/\sqrt{2})$ verteilt ist:

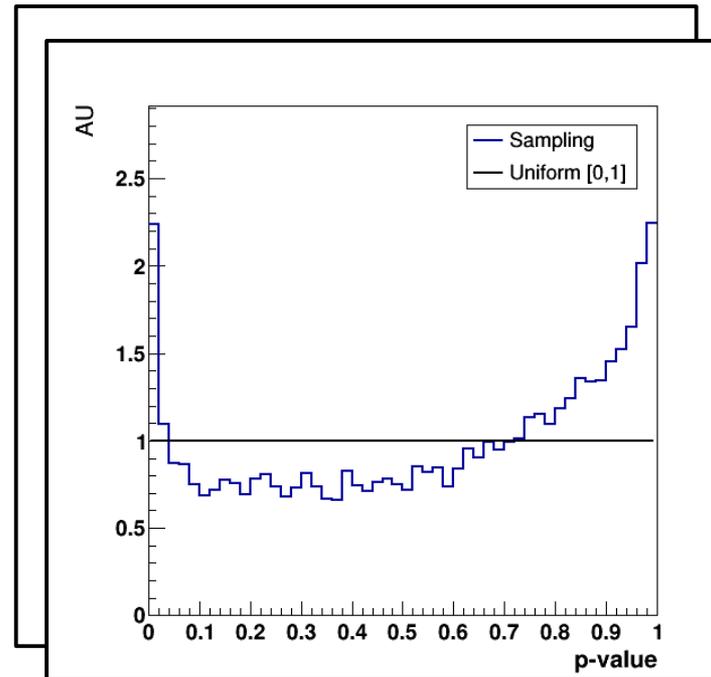
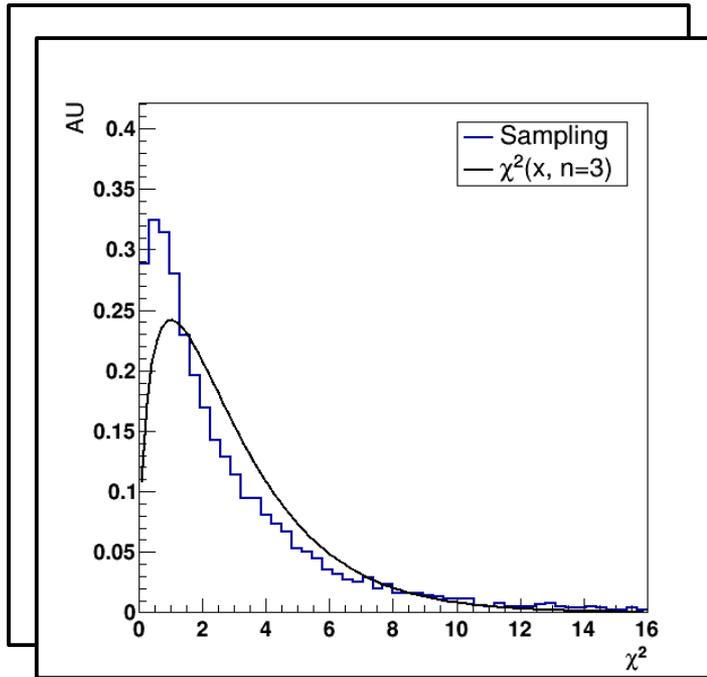


Ein lauffähiges ROOT macro finden Sie [hier](#).

- Die Verteilungen der LS-Schätzungen $\hat{\theta}_0$ und $\hat{\theta}_1$ sind jeweils normalverteilt und unverzerrt, aber nicht mehr effizient.

Verteilungsfreiheit – zugrundeliegende exp Verteilung

- Hier sehen Sie die Ausgänge der Messung nach 10k-facher Wiederholung des Pseudo-experiments, wobei jede Einzelmessung beidseitig nach $\exp(x, \sigma_i/\sqrt{2})$ verteilt ist:



Verteilungsfreiheit – Zusammenfassung

- Unabhängig von der zugrundeliegenden Wahrscheinlichkeitsdichte für jede Einzelmessung sind die LS-Schätzwerte normalverteilt und für symmetrische Wahrscheinlichkeitsdichten erwartungstreu!
- Wie sich verschiedene zugrundeliegende Wahrscheinlichkeitsdichten äußern:
 - Die LS-Schätzung ist nicht mehr effizient (vgl. [Lecture_05](#) Folie 23).
 - χ_{obs}^2 ist nicht (mehr) nach $\chi^2(x, n)$ verteilt.
 - Der aus dem Integral der $\chi^2(x, n)$ Verteilung bestimmte p-Wert ist nicht mehr uniform verteilt, wie dies der Fall wäre, wenn die LS-Schätzung einer $\chi^2(x, n)$ Verteilung folgen würde.

pull-Verteilung

Für normalverteilte Messwerte sollte jeder LS-Schätzwert nach Standardisierung

$$Z_j = \frac{\hat{\theta}_j - E[\hat{\theta}_j]}{\sqrt{\text{var}[\hat{\theta}_j]}} = \frac{\hat{\theta}_j - \theta_j}{\sqrt{\text{var}[\hat{\theta}_j]}}$$

einer Standardnormalverteilung $\varphi(x, 0, 1)$ folgen.

- Man nennt die Verteilung der Werte für Z_j nach mehrfacher Durchführung von Pseudoexperimenten *pull*-Verteilung. Dabei werden bei bekannt vorausgesetzten Werten für θ_j die Werte von $\hat{\theta}_j$ und

$$\sqrt{\text{var}[\hat{\theta}_j]}$$

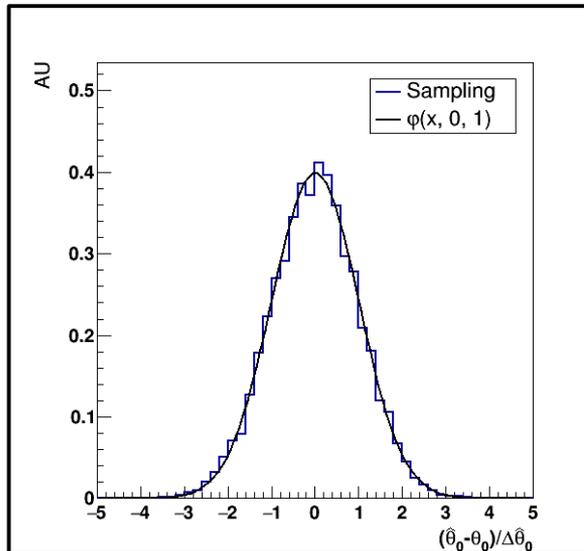
für jeden Ausgang des Pseudoexperiments aus dem LS-Schätzwert und seiner Varianz bestimmt.

- Aus der *pull*-Verteilung wird z.B. üblicherweise die Erwartungstreue der Abschätzung bestimmt.

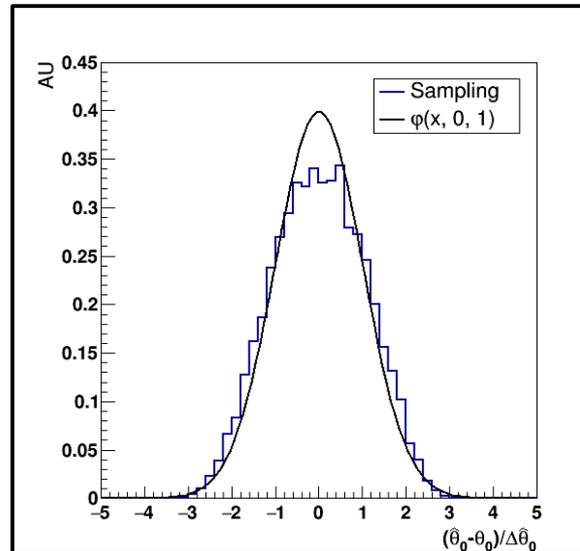
pull-Verteilung – am Beispiel

- Wir zeigen die *pull*-Verteilung wieder anhand des Beispiels aus [Lecture_05](#) (Folie 25, „Anpassung einer Geraden an fünf Messpunkte“):

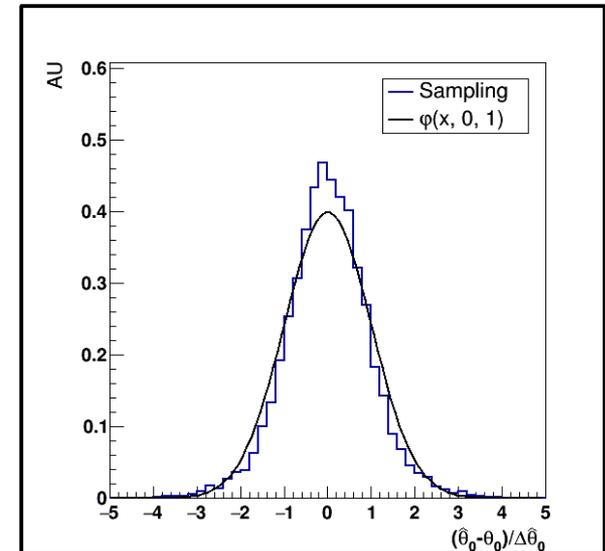
Messwerte normalverteilt.



Messwerte gleichverteilt.



Messwerte doppelt
exponentiell verteilt.



Ein lauffähiges ROOT macro finden Sie [hier](#).

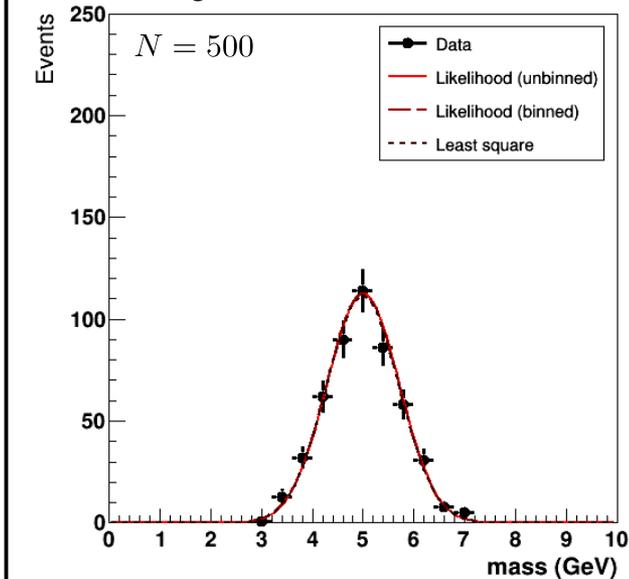
- Die LS-Schätzwerte sind zwar in allen Fällen erwartungstreu, ihre *pull*-Verteilungen sind jedoch nur im linken Fall standardnormalverteilt.

LS-Schätzung bei schwach populierten Histogramm-Bins

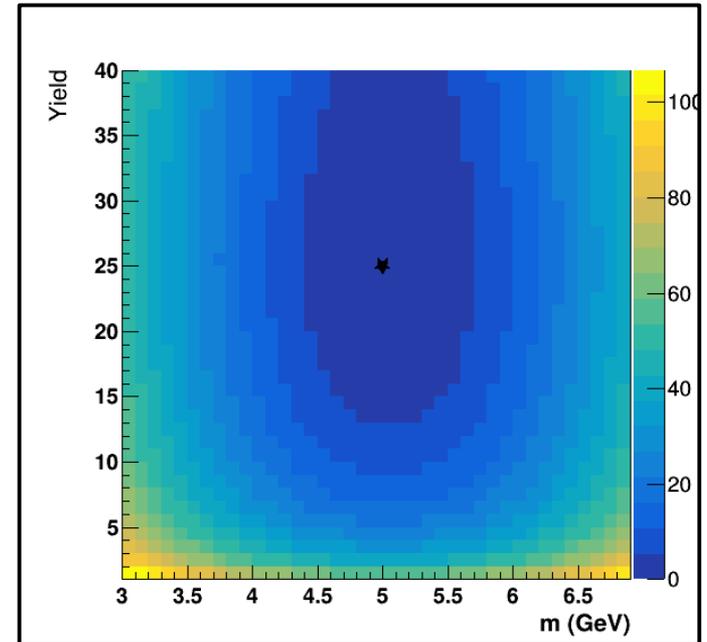
- Im Fall leerer oder schwach populierter Bins ist die LS-Schätzung, im Gegensatz zur ML-Schätzung, nicht mehr erwartungstreu. Wir demonstrieren dies anhand des folgenden Beispiels aus der Teilchenphysik:

Resonanz ohne Untergrund nach $\varphi(x, \mu = 5, \sigma = 1)$ verteilt. **N Ereignisse beobachtet.**

Abschätzung



Ein lauffähiges RooT macro finden Sie [hier](#).



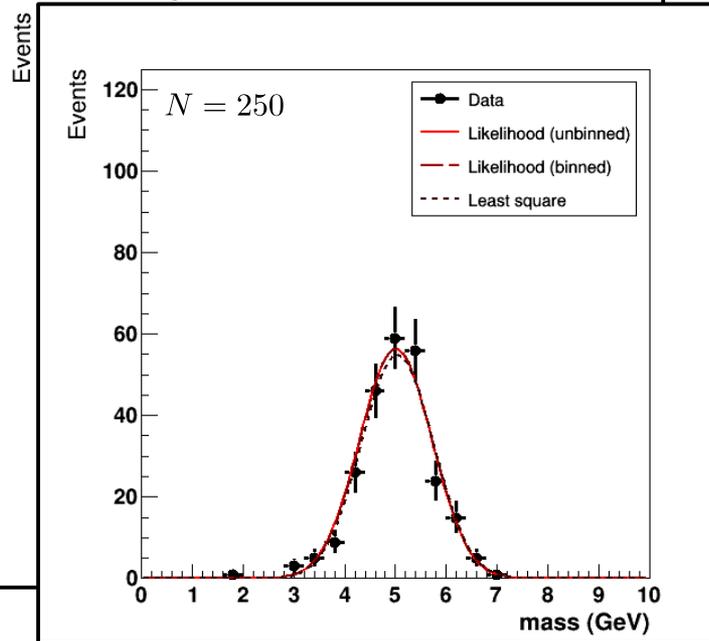
Ungebinteter Likelihood scan

LS-Schätzung bei schwach populierten Histogramm-Bins

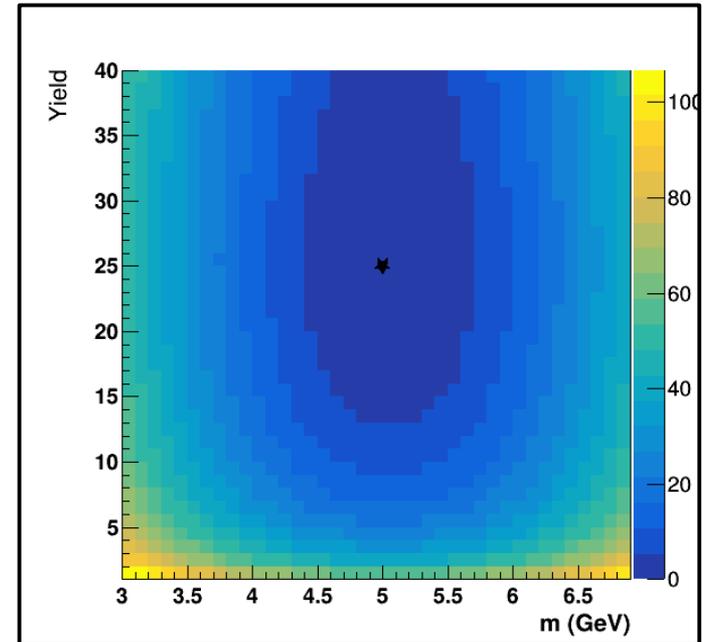
- Im Fall leerer oder schwach populierter Bins ist die LS-Schätzung, im Gegensatz zur ML-Schätzung, nicht mehr erwartungstreu. Wir demonstrieren dies anhand des folgenden Beispiels aus der Teilchenphysik:

Resonanz ohne Untergrund nach $\varphi(x, \mu = 5, \sigma = 1)$ verteilt. **N Ereignisse beobachtet.**

Abschätzung



Ein lauffähiges RooT macro finden Sie [hier](#).



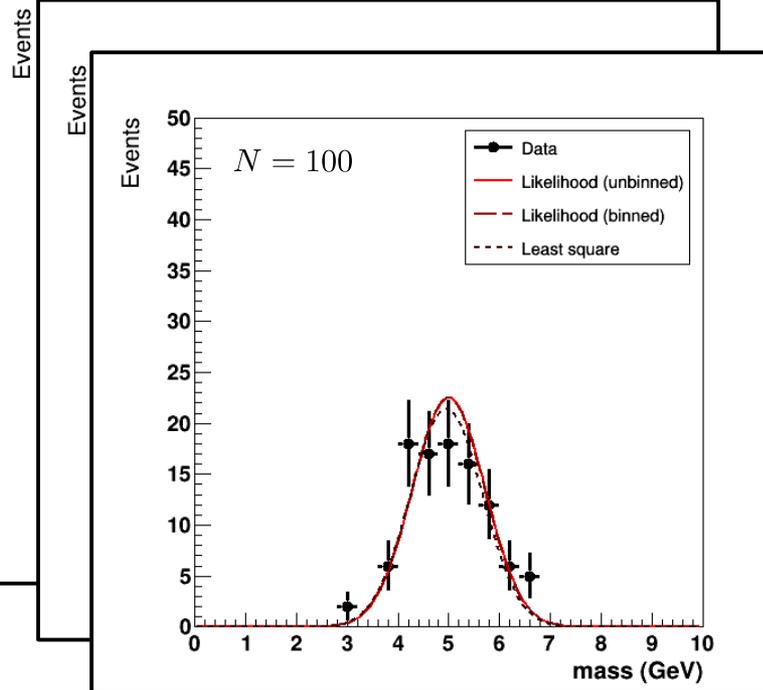
Ungebintter Likelihood scan

LS-Schätzung bei schwach populierten Histogramm-Bins

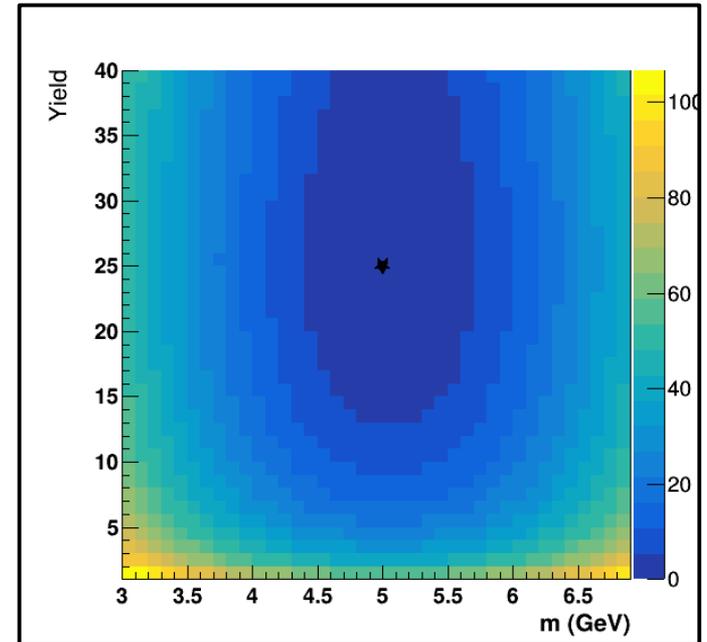
- Im Fall leerer oder schwach populierter Bins ist die LS-Schätzung, im Gegensatz zur ML-Schätzung, nicht mehr erwartungstreu. Wir demonstrieren dies anhand des folgenden Beispiels aus der Teilchenphysik:

Resonanz ohne Untergrund nach $\varphi(x, \mu = 5, \sigma = 1)$ verteilt. **N Ereignisse beobachtet.**

Abschätzung



Ein lauffähiges RooT macro finden Sie [hier](#).



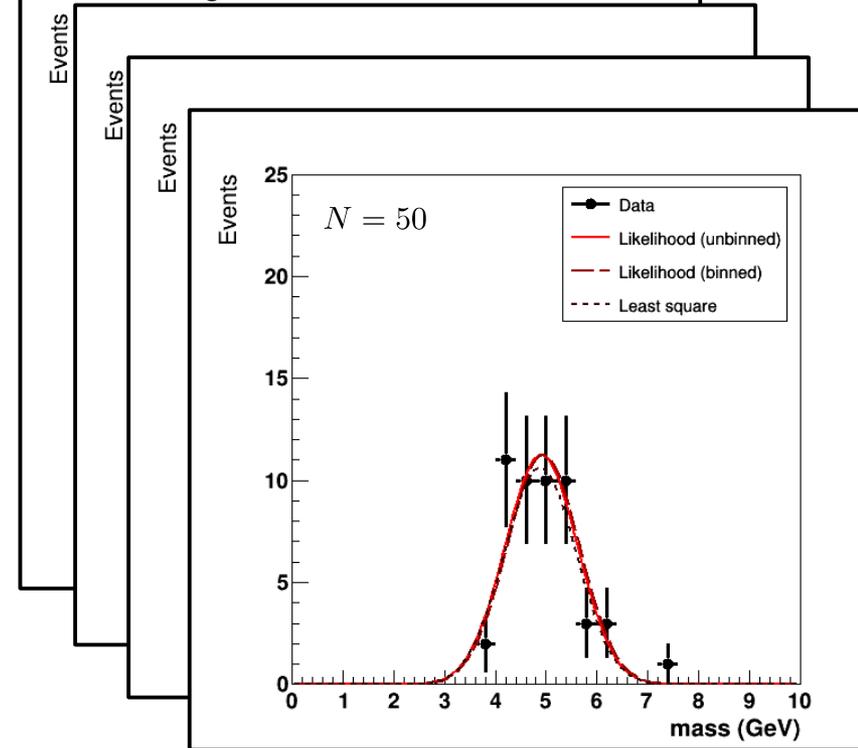
Ungebintter Likelihood scan

LS-Schätzung bei schwach populierten Histogramm-Bins

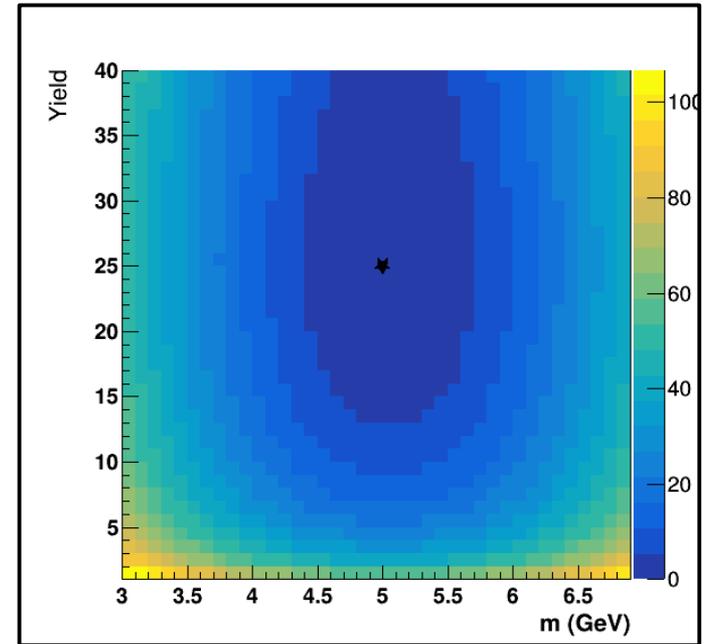
- Im Fall leerer oder schwach populierter Bins ist die LS-Schätzung, im Gegensatz zur ML-Schätzung, nicht mehr erwartungstreu. Wir demonstrieren dies anhand des folgenden Beispiels aus der Teilchenphysik:

Resonanz ohne Untergrund nach $\varphi(x, \mu = 5, \sigma = 1)$ verteilt. **N Ereignisse beobachtet.**

Abschätzung



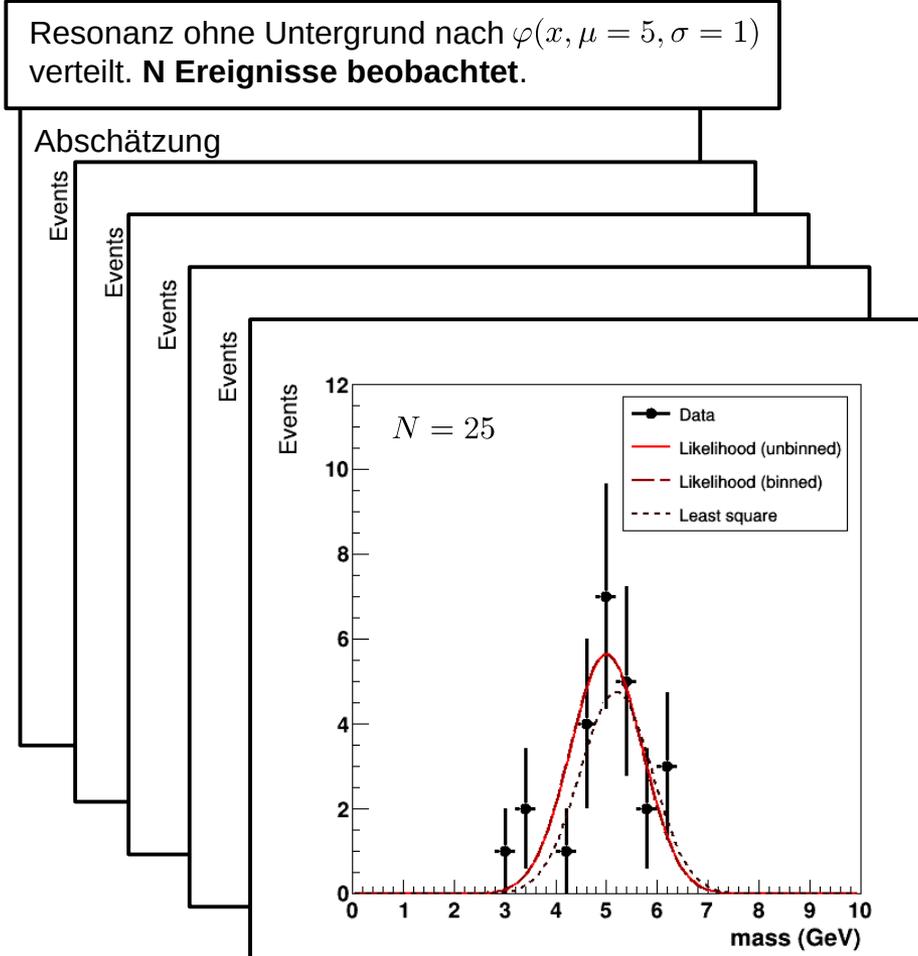
Ein lauffähiges RooT macro finden Sie [hier](#).



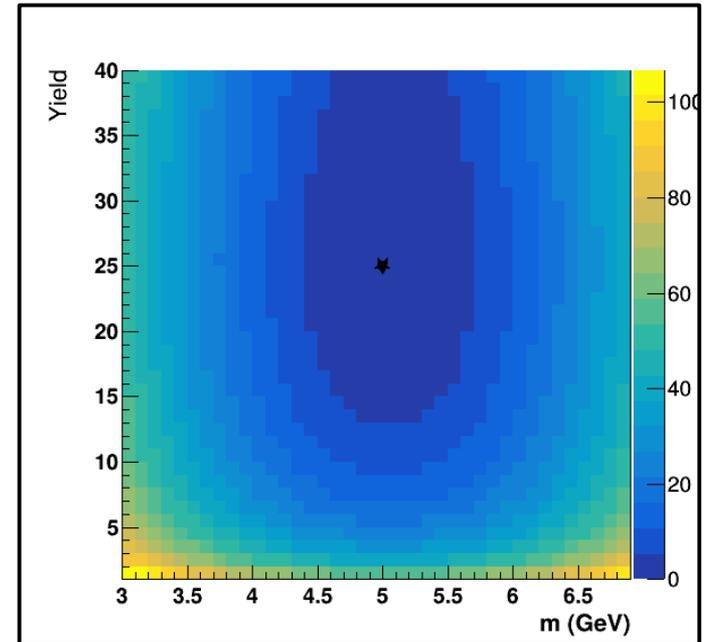
Ungebintter Likelihood scan

LS-Schätzung bei schwach populierte Histogramm-Bins

- Im Fall leerer oder schwach populierter Bins ist die LS-Schätzung, im Gegensatz zur ML-Schätzung, nicht mehr erwartungstreu. Wir demonstrieren dies anhand des folgenden Beispiels aus der Teilchenphysik:



Ein lauffähiges RooT macro finden Sie [hier](#).



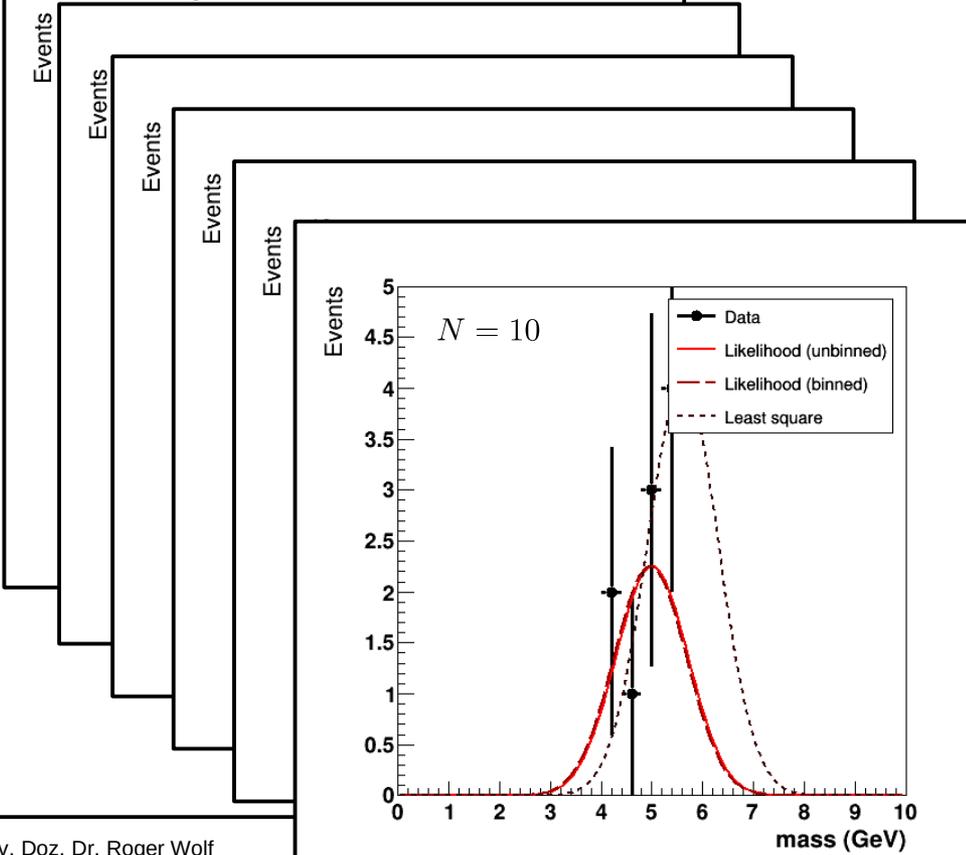
Ungebintter Likelihood scan

LS-Schätzung bei schwach populierten Histogramm-Bins

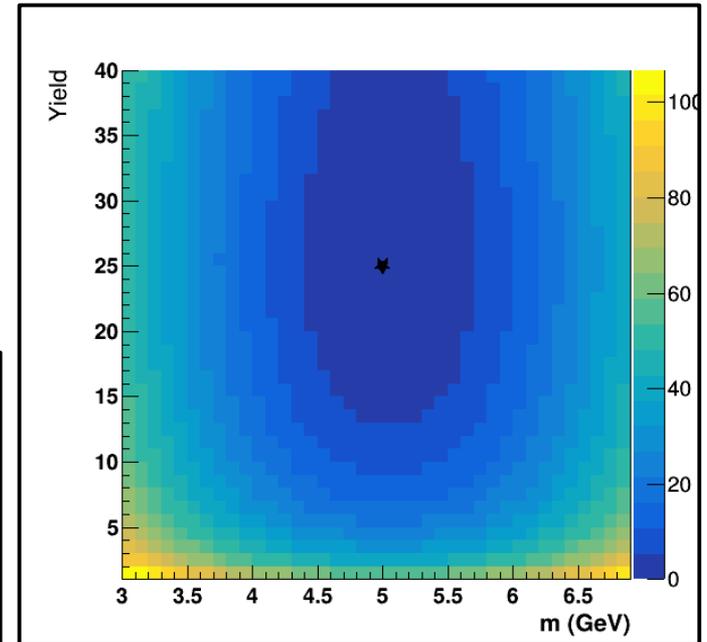
- Im Fall leerer oder schwach populierter Bins ist die LS-Schätzung, im Gegensatz zur ML-Schätzung, nicht mehr erwartungstreu. Wir demonstrieren dies anhand des folgenden Beispiels aus der Teilchenphysik:

Resonanz ohne Untergrund nach $\varphi(x, \mu = 5, \sigma = 1)$ verteilt. **N Ereignisse** beobachtet.

Abschätzung



Ein lauffähiges RooT macro finden Sie [hier](#).



Ungebinteter Likelihood scan

LS- vs. ML-Schätzung

- Nutzen Sie die **ML-Schätzung** wann immer möglich:
 - Sie setzt die Hypothese einer parametrisierten zugrundeliegende Wahrscheinlichkeitsdichte voraus. Die Parameter werden maximiert.
 - Sie ist immer erwartungstreu.
 - Wenn es eine effiziente Schätzung des gesuchten Parameters gibt ist die ML-Schätzung effizient, d.h. sie hat die geringste Varianz.
 - Die ML-Schätzung ist i.a. nicht analytisch lösbar. Heutzutage würden Sie in ernsthaften Studien zur Lösung statistischer Probleme ohnehin auf MC Methoden zurückgreifen.

LS- vs. ML-Schätzung

- Die **LS-Schätzung** benötigt weniger Voraussetzungen von Ihrer Seite: Nur der Erwartungswert und die Varianz der Hypothese müssen bekannt sein, die hypothetische zugrundeliegende Wahrscheinlichkeitsdichte ist die Normalverteilung:
 - Die LS-Schätzung bezieht ihre große Bedeutung daraus, dass sie für lineare Probleme analytisch lösbar ist.
 - Für eine größere Klasse an wohldefinierten Problemen ist die Lösung (z.B. F-Test und T-Test) ist die Lösung bekannt.
 - Für normalverteilte Zufallsgrößen ist die LS-Schätzung zur ML-Schätzung äquivalent.
 - Für nicht-normalverteilte Zufallsgrößen und insbesondere für asymmetrische Unsicherheiten führen die LS- und die ML-Schätzung nicht zum gleichen Ergebnis.

Goodness-of-fit Tests

Goodness-of-Fit Test

- Mit Hilfe des Goodness-of-fit (GoF) Tests überprüfen Sie ob und wieweit Ihr Modell Ihre Daten/Messung beschreiben kann.

Der Goodness-of-fit Test (GoF) basiert auf der Schätzfunktion $t_{\text{GoF}}(\vec{x})$. Die Einschätzung des Tests erfolgt auf Grundlage des p -Wertes (p -value) der Stichprobenverteilung $g(t_{\text{GoF}}(\vec{x}))$

$$p = \int_{t_{\text{GoF}}^{\text{obs}}}^{\infty} g(t'_{\text{GoF}}) dt'_{\text{GoF}} \quad \text{mit } 0 \leq t_{\text{GoF}}(\vec{x}) < \infty$$

mit der Interpretation:

t_{GoF} klein: Das Modell kann die Daten gut beschreiben

t_{GoF} groß: Das Modell kann die Daten nicht gut beschreiben

Der p -Wert bezeichnet die Wahrscheinlichkeit einen Wert von $t_{\text{GoF}} \geq t_{\text{GoF}}^{\text{obs}}$ zu erhalten, wenn das Modell wahr ist.

- Der GoF ist ein spezieller Hypothesentest.

Vorgehensweise beim GoF Test

- Die Berechnung des p-Wertes impliziert folgendes Vorgehen:
 - (Einmalige) Evaluation von $t_{\text{GoF}}^{\text{obs}}$ auf Daten.
 - Vielfache Auswertung von t_{GoF} auf randomisierten Daten mit Hilfe der MC Methode zur Bestimmung von $g(t_{\text{GoF}}(\vec{x}))$.
 - Bestimmung von p .
- Diese Vorgehensweise ist die beste und sicherste, um ein Modell gegen eine Messung zu testen.
- In seltenen Fällen ist $g(t_{\text{GoF}}(\vec{x}))$ analytisch bekannt, so z.B. beim χ^2 -Test nach Pearson. In diesen Fällen kann die Einschätzung auch direkt auf Grundlage von $t_{\text{GoF}}^{\text{obs}}$ erfolgen.

Pearson (χ^2)-Test

Die Pearson Teststatistik

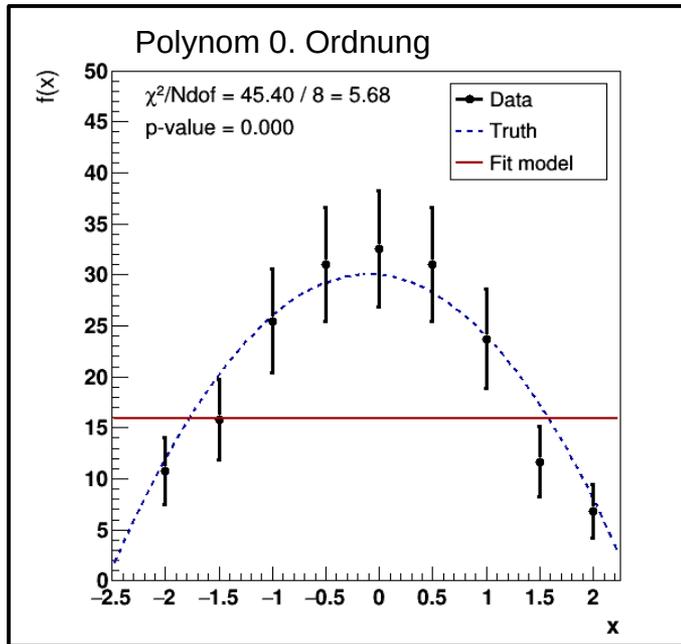
$$t_{\chi^2} = \sum_{i \leq n} \frac{(f(x_i) - y_i)^2}{f(x_i)}$$

entspricht dem χ^2 -Wert für Poisson-verteilte Zufallsgrößen mit $\sigma_i^2 = f(x_i)$.

- Die $f(x_i)$ entsprechen den Modellvorhersagen, die y_i den Messungen/Beobachtungen.
- t_{χ^2} folgt einer χ^2 Verteilung, d.h. $\frac{t_{\chi^2}}{n} \lesssim 1$ impliziert, dass das Modell die Daten beschreiben kann.
- **NB:** Das ist nicht gleichbedeutend damit, dass das Modell der Wahrheit entspricht.
- Der Pearson Test kann einfach und schnell bei vielen Gelegenheiten angewandt werden.

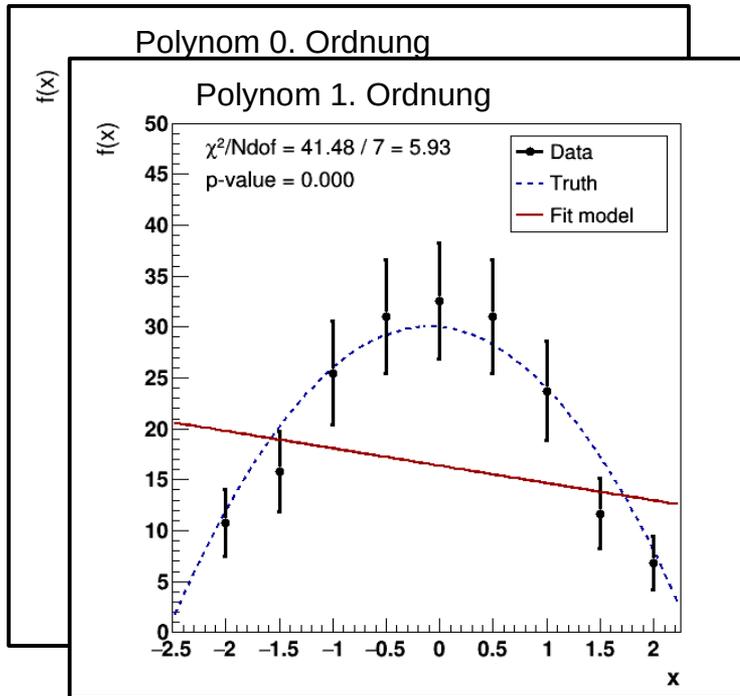
Beispiel – Anpassung eines Polynoms

- Anpassung eines Polynoms variierender Ordnung an eine randomisierte Verteilung mit dem zugrundeliegenden Verlauf $f(x) = 30 - x - 5x^2$:



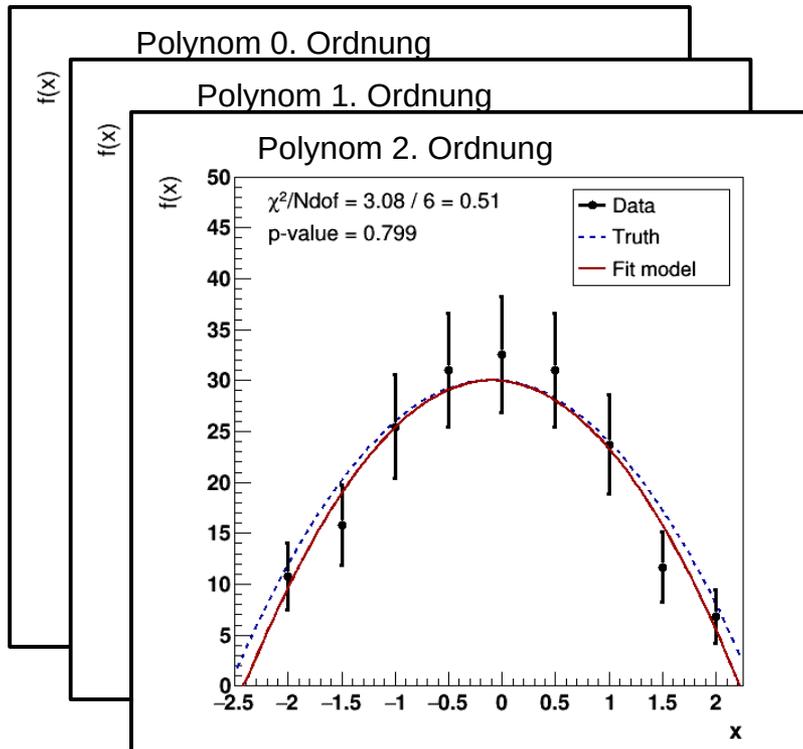
Beispiel – Anpassung eines Polynoms

- Anpassung eines Polynoms variierender Ordnung an eine randomisierte Verteilung mit dem zugrundeliegenden Verlauf $f(x) = 30 - x - 5x^2$:



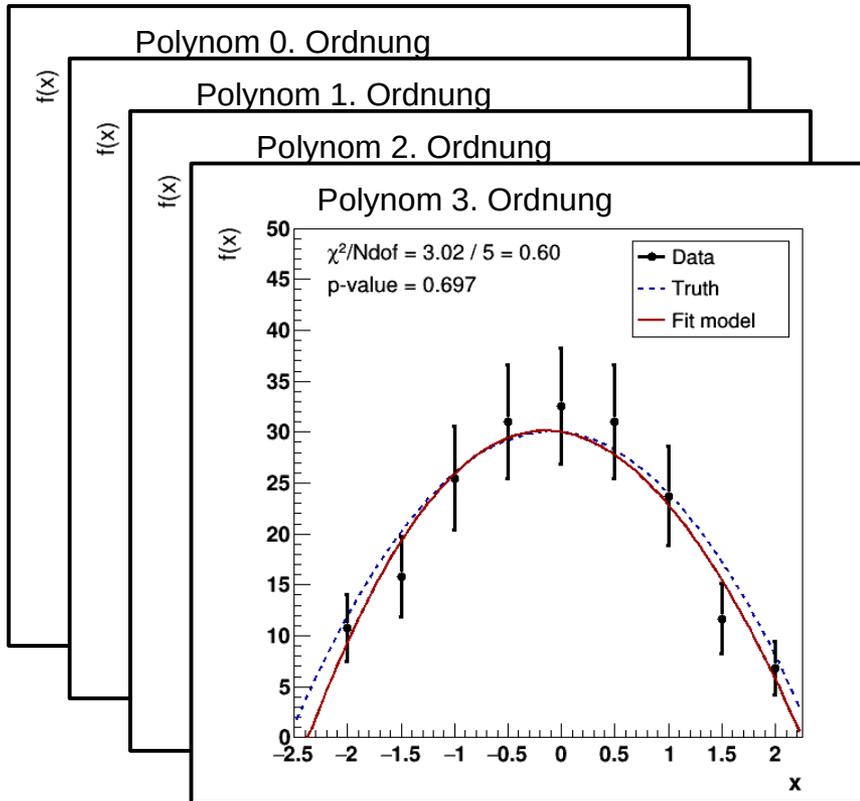
Beispiel – Anpassung eines Polynoms

- Anpassung eines Polynoms variierender Ordnung an eine randomisierte Verteilung mit dem zugrundeliegenden Verlauf $f(x) = 30 - x - 5x^2$:



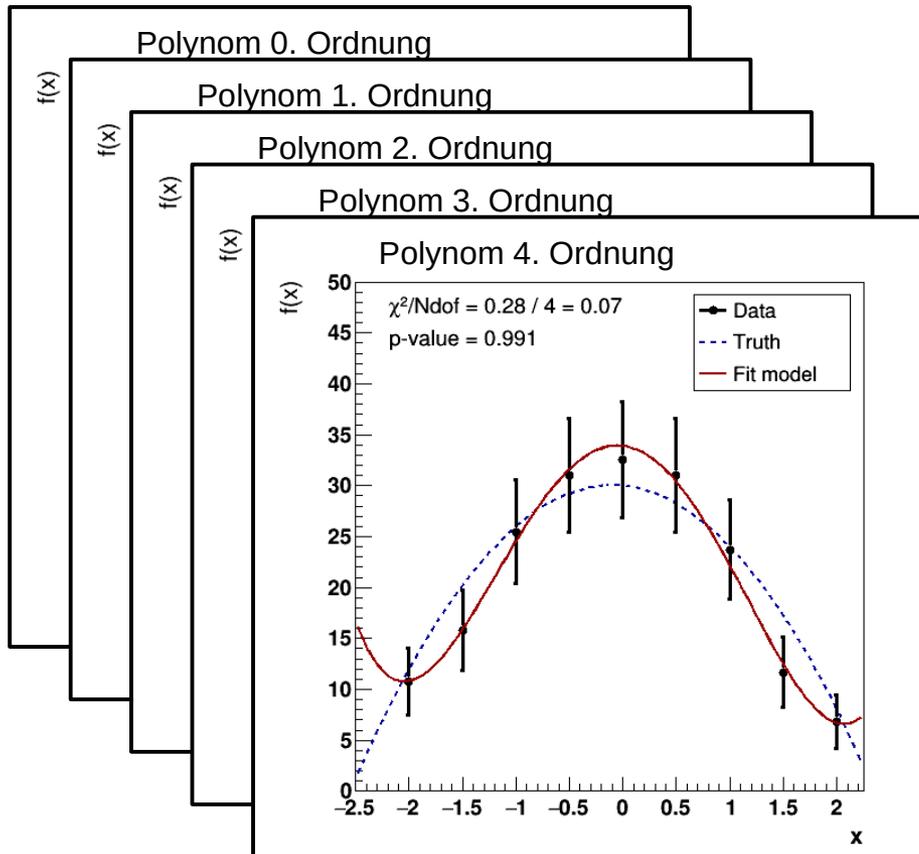
Beispiel – Anpassung eines Polynoms

- Anpassung eines Polynoms variierender Ordnung an eine randomisierte Verteilung mit dem zugrundeliegenden Verlauf $f(x) = 30 - x - 5x^2$:



Beispiel – Anpassung eines Polynoms

- Anpassung eines Polynoms variierender Ordnung an eine randomisierte Verteilung mit dem zugrundeliegenden Verlauf $f(x) = 30 - x - 5x^2$:



Ein lauffähiges Root macro finden Sie [hier](#).

- Halten Sie bei Anpassungen die Anzahl nicht a priori motivierter Parameter so gering wie möglich.
- Werte von $\chi^2 \ll 1$ können nicht nur auf überschätzte Unsicherheiten sondern auch auf „*overfitting*“ hindeuten.
- Beachten Sie: Bei guter Abdeckung Ihrer Unsicherheiten berühren ~30% der Datenpunkte die angepasste Kurve i.a. nicht mit ihren Fehlerbalken.

Likelihood Quotient

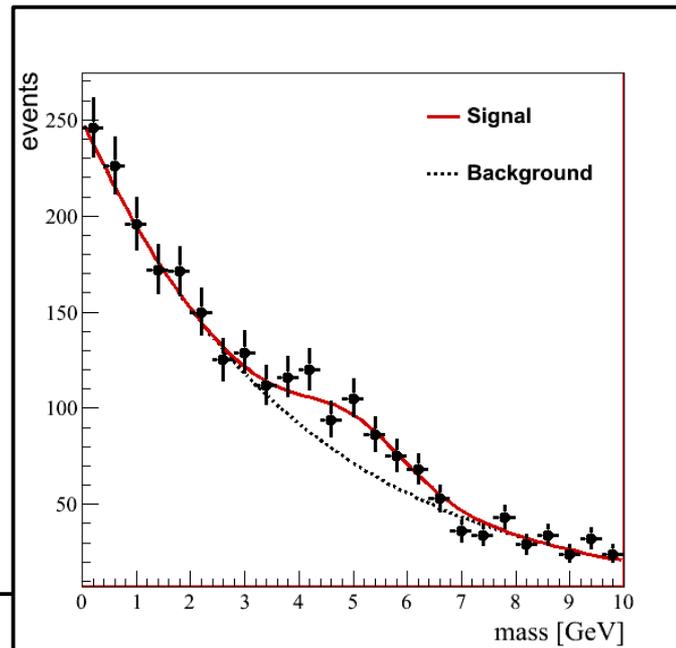
- Der sogenannte saturierte Modelltest erweitert den χ^2 -Test nach Pearson zu einem allgemeinen Likelihood basierten Test.
- Die Teststatistik ist der Likelihood Quotient:

$$t_{\text{sat.}} = -2 \ln \left(\frac{\mathcal{L}(\text{data} |_{\text{test}})}{\mathcal{L}(\text{data} |_{\text{sat.}})} \right)$$

Zu testendes
Modell

Modell mit ebenso vielen
Parametern, wie Messungen
(sog. Saturiertes Modell)

Im rechten Bsp. hätte das
saturierte Modell 25 Para-
meter und würde so durch
jeden Messpunkt gehen.



Likelihood Quotient

- Der sogenannte saturierte Modelltest erweitert den χ^2 -Test nach Pearson zu einem allgemeinen Likelihood basierten Test.
- Die Teststatistik ist der Likelihood Quotient:

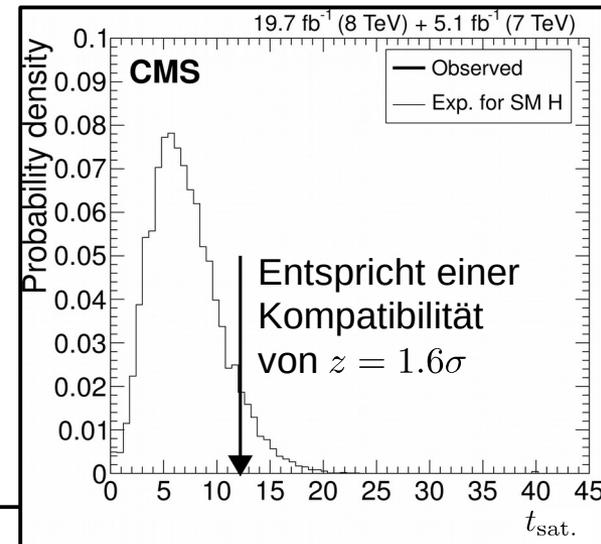
$$t_{\text{sat.}} = -2 \ln \left(\frac{\mathcal{L}(\text{data}|_{\text{test}})}{\mathcal{L}(\text{data}|_{\text{sat.}})} \right)$$

- Für Histogramme ohne weitere systematischen Unsicherheiten, für die die Unsicherheiten der Messpunkte normalverteilt sind erhält man für große Zählstatistik $t_{\text{sat.}} \rightarrow t_{\chi^2}$ zurück:

$$\mathcal{L}(\text{data}|_{\text{test}}) = \prod_{i \leq n} \frac{1}{\sqrt{2\pi\sigma_i}} e^{-(y_i - \mu_i)^2/2\sigma_i}$$

$$\mathcal{L}(\text{data}|_{\text{sat.}}) = \prod_{i \leq n} \frac{1}{\sqrt{2\pi\sigma_i}}$$

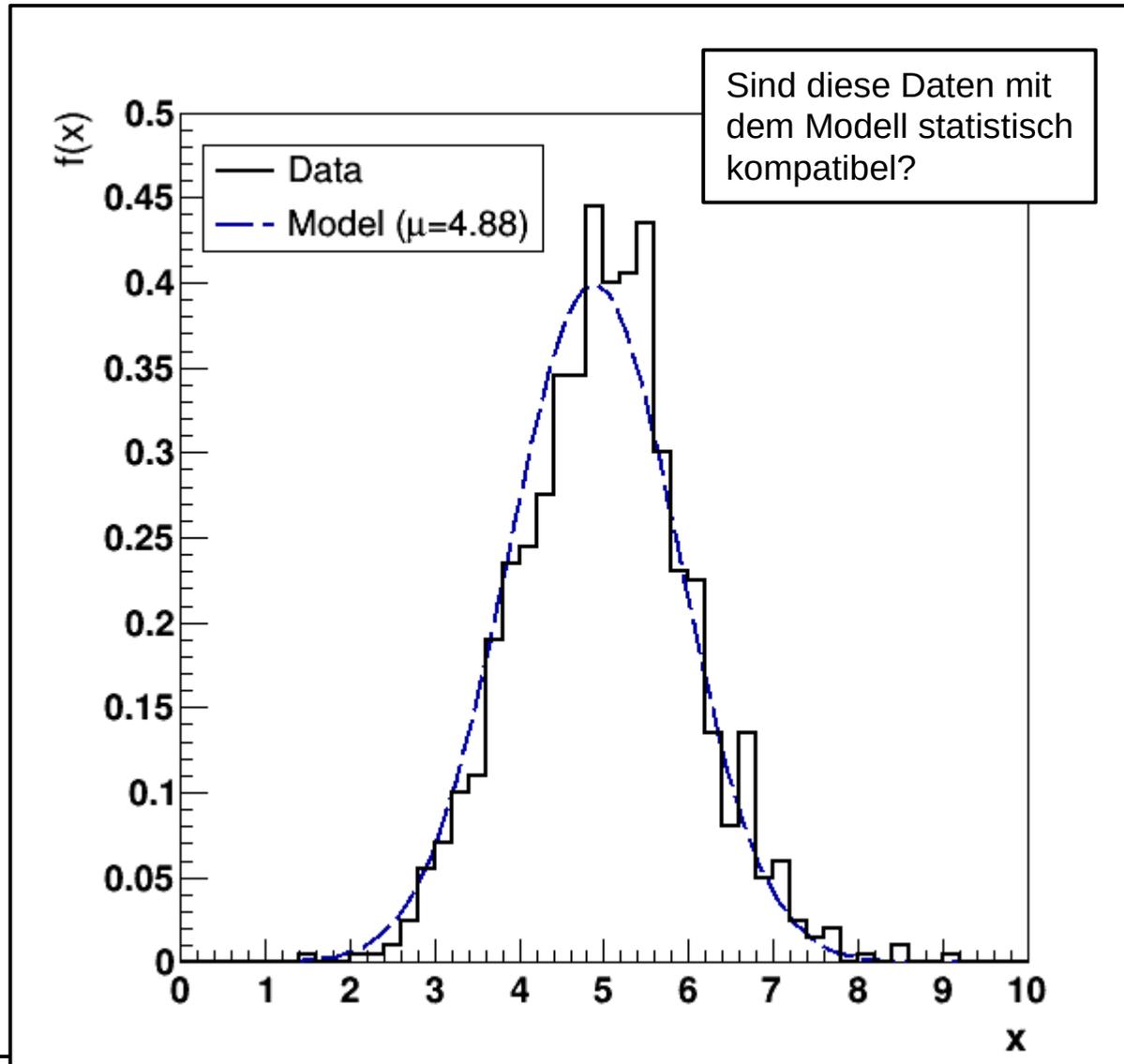
$$t_{\text{sat.}} = -2 \ln \left(\frac{\mathcal{L}(\text{data}|_{\text{test}})}{\mathcal{L}(\text{data}|_{\text{sat.}})} \right) = \sum_{i \leq n} \frac{(y_i - \mu_i)^2}{\sigma_i}$$



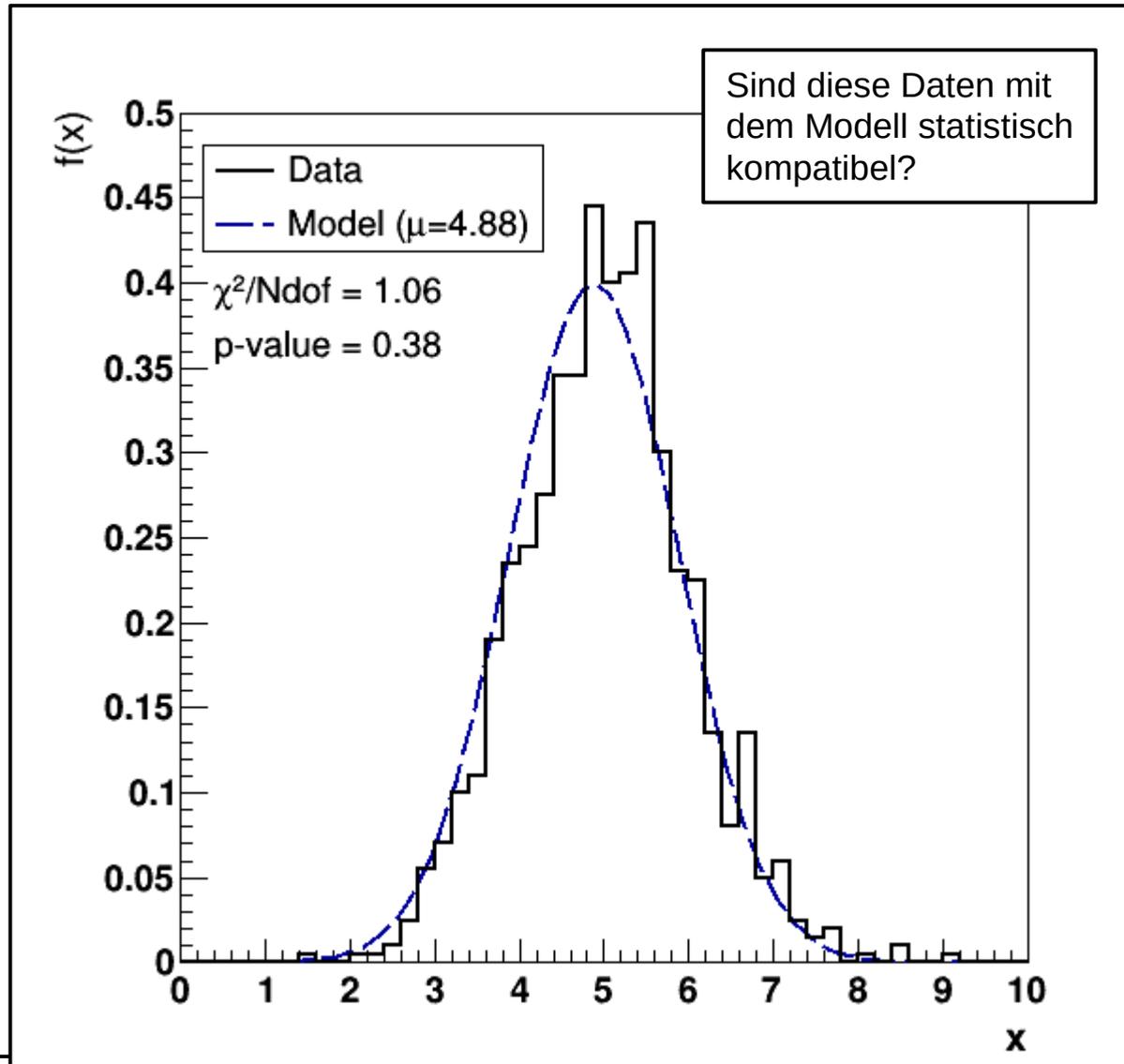
Tests auf der CDF

- Die Teststatistik t_{χ^2} ist nur sensitiv auf die Abstände der Erwartung von den Messpunkten, nicht auf deren „Richtung“. Warum ist dies so?
- Es kann somit Vorkommen, dass Sie bin-korrelierte Trends nicht mit einem auf t_{χ^2} oder $t_{\text{sat.}}$ basierten Test erfassen.
- Dies können Sie mit Tests erreichen, die auf die kummulative (CDF) und die empirische (EDF) Verteilungsfunktion sensitiv sind. Wir geben als Beispiel eines solchen Tests den [Kolmogorov-Smirnov-Test](#) (KS) an:

Vergleich zweier Verteilungsfunktionen

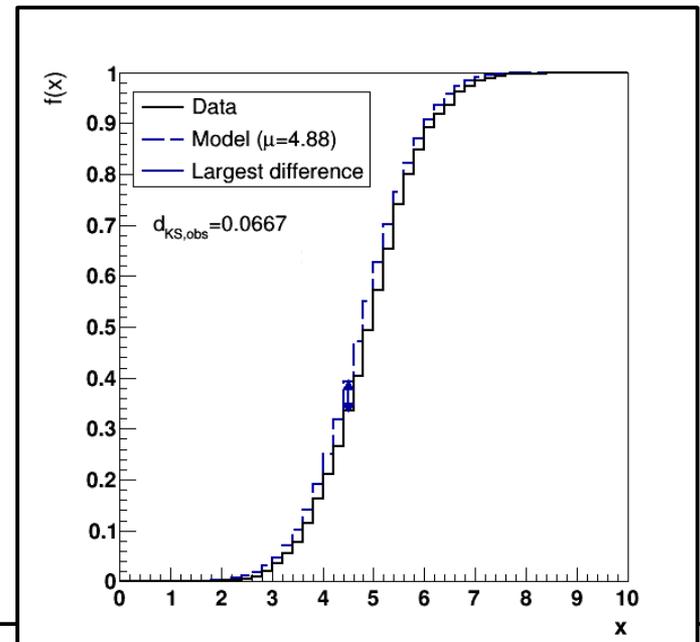
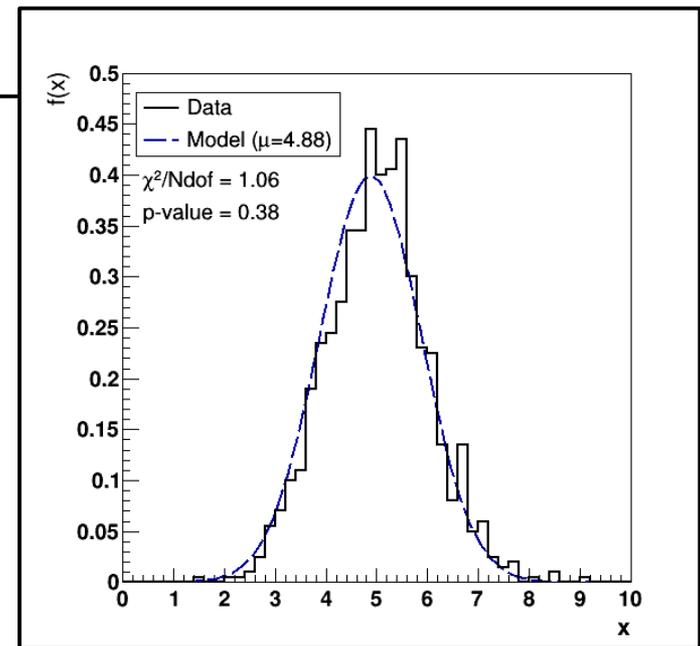


Vergleich zweier Verteilungsfunktionen



Stat. Kompabilität und Trends

- Für dieses Bsp. wurden die Daten basierend auf einer Normalverteilung mit $\mu = 5$ 1000 mal zufällig gesampled. Der Vergleich erfolgt mit einem Modell mit $\mu = 4.88$.
- Wie Sie sehen ist das Modell mit $\mu = 4.88$ nach χ^2 mit den Daten kompatibel. Es liegt jedoch, wenn Sie genau hinsehen, auf der linken Flanke systematisch unter- und auf der rechten Flanke oberhalb der Daten.
- Diesen Trend erfassen Sie durch den Vergleich der CDF des Modells mit der EDF der Daten.



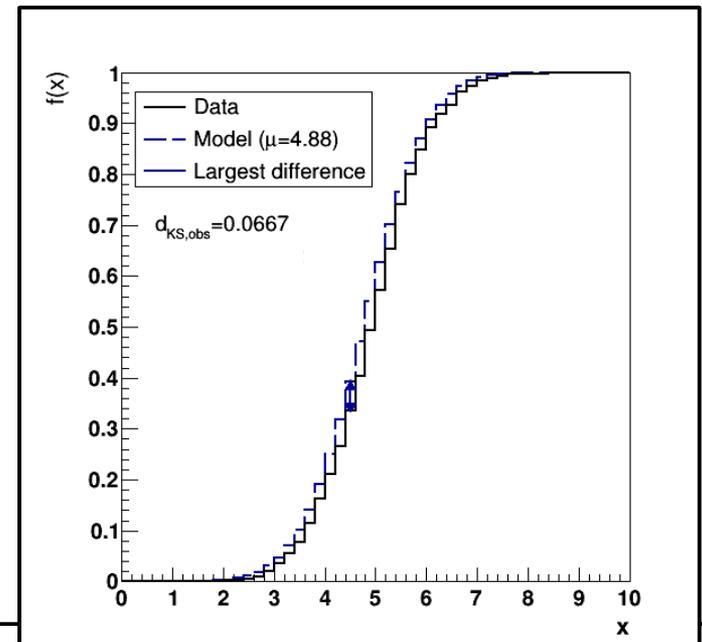
Kolmogorov-Smirnov-Test

Der größte vertikale Abstand zweier kummulativer Verteilungsfunktionen F_0 und F_n

$$d_{KS} = \|F_n - F_0\|$$

ist die Teststatistik des Kolmogorov-Smirnov-Tests.

- Diesen Trend erfassen Sie durch den Vergleich der CDF des Modells mit der EDF der Daten.
- Die Werte d_{KS} können Sie aus [Tabellen](#) auslesen.



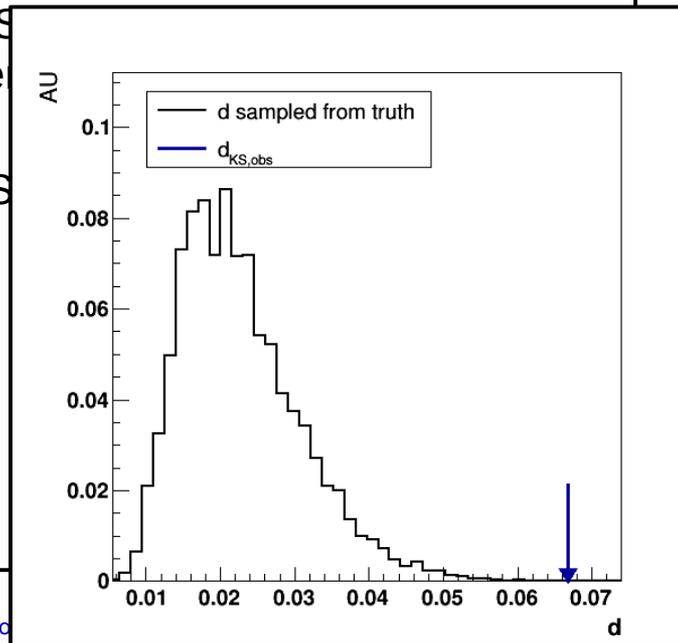
Kolmogorov-Smirnov-Test

Der größte vertikale Abstand zweier kummulativer Verteilungsfunktionen F_0 und F_n

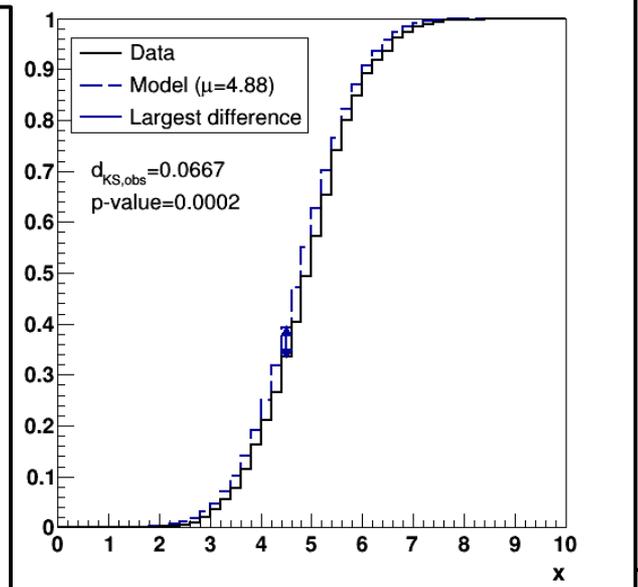
$$d_{KS} = \|F_n - F_0\|$$

ist die Teststatistik des Kolmogorov-Smirnov-Tests.

- Diesen Trend erfassen S
CDF des Modells mit de
- Die Werte d_{KS} können S
- Die sicherste Art einen p-Wert für einen beobachteten Wert $d_{KS,obs}$ zu ermitteln besteht in der MC Methode.



Ein lauffähiges Root macro finden Sie [hier](#).



Zusammenfassung

- LS-Schätzung:
 - Bedeutung von χ_{obs}^2 .
 - Verteilungsfreiheit.
 - *pull*-Verteilung.
 - LS-schätzung bei schwach populierte Histogramm-Bins.
- Goodness-of-fit Tests:
 - Pearson-Test.
 - Saturierter Modelltest.
 - Kolmogorov-Smirnov-Test.
 - Ein alltäglicheres Beispiel für den diskutierten Pearson-Test finden Sie im Backup. Beispiele weiterer Tests finden sich in der Vorlesung von Quast/Meyer.
 - Wenn Sie die diskutierten GoF Tests an Beispielen aus dem Leben diskutiert sehen möchten werfen Sie einen Blick in die [Masterarbeit von Dennis Roy](#).

Backup

Beispiel: Test auf Unabhängigkeit

- Zu testende **Hypothese**: Bei einer Umfrage zur Rentenreform soll in Abhängigkeit vom Alter die Zustimmung (-1 ablehnend, 0 neutral +1 zustimmend) erfasst werden. Man vermutet, dass das Alter einen Einfluss auf die Zustimmung hat.

- H_0 : „Die Häufigkeiten in der folgenden Kontingenztabelle der Grundgesamtheiten unterscheiden sich nicht von der Erwartung für statistische Unabhängigkeit“
($H_0 : p_{ij} = p_{ij}^e$).

- Testfunktion:

$$t_{\chi^2} = \sum_{i=1}^m \sum_{j=1}^k \frac{h_{ij} - h_{ij}^e}{h_{ij}^e} \propto \chi_{\alpha}^2 ((m-1) \cdot (k-1)) = \chi_{\alpha}^2(6)$$

- Irrtumswahrscheinlichkeit: $\alpha = 5\%$ mit: $(\chi_{0.05}^2(6) = 12,6)$

Beispiel: Test auf Unabhängigkeit

- Stichprobenergebnis (Befragung von 250 Personen):

Einstellung	Alter			
	- 20	21-40	41-65	66 und älter
-1	5	30	60	30
0	6	24	30	15
+1	9	26	10	5

$$\sum \chi^2 =$$

		Alter			
Einstellung		-20	21-40	41-65	66 und älter
h_{ij}	h_{ij}^e				
$(h_{ij} - h_{ij}^e)^2$	$\frac{(h_{ij} - h_{ij}^e)^2}{h_{ij}^e}$				
	-1				
	0				
	+1				

Beispiel: Test auf Unabhängigkeit

- Stichprobenergebnis (Befragung von 250 Personen):

Einstellung	Alter			
	- 20	21-40	41-65	66 und älter
-1	5	30	60	30
0	6	24	30	15
+1	9	26	10	5

$$\sum \chi^2 = 28$$

h_{ij}	h_{ij}^e			Einstellung				Alter					
				-20	21-40	41-65	66 und älter						
$(h_{ij} - h_{ij}^e)^2$	$\frac{(h_{ij} - h_{ij}^e)^2}{h_{ij}^e}$			5	10	30	40	60	50	30	25		
		-1		25	2,5	100	2,5	100	2	25	1		
		0		6	6	24	24	30	30	15	15		
				0	0	0	0	0	0	0	0		
		+1		9	4	26	16	10	20	5	10		
				25	6,25	100	6,25	100	5	25	2,5		

Beispiel: Test auf Unabhängigkeit

- Testentscheidung: H_0 ist widerlegt ($\alpha^* = 0$), es besteht also eine Korrelation zwischen Alter und Zustimmung zur Rentenreform. Welcher Art diese Korrelation ist und wie sie interpretierbar ist sollte Bestandteil weiterer Untersuchungen sein.