Statistische Methoden der Datenanalyse

Prof. Markus Schumacher

ALU Freiburg, Wintersemsester 2009/2010

BOK-Veranstaltung im Rahmen des ZfS

- Kapitel 4: Grundlagen der Parameterschätzung

Schätzer für den Mittelwert der Grundgesamtheit μ

$\widehat{\mu} = \overline{x} = \frac{1}{n}\sum_{i=1}^{n} x_i$ den Mittelwert der Stichprobe

$\widehat{\mu} = \frac{1}{10}\sum_{i=1}^{10} x_i$ den Mittelwert der ersten 10 Punkte der Stichprobe

$\widehat{\mu} = \frac{1}{n-1}\sum_{i=1}^{n} x_i$ $n/(n-1)$ mal den Mittelwert der Stichprobe

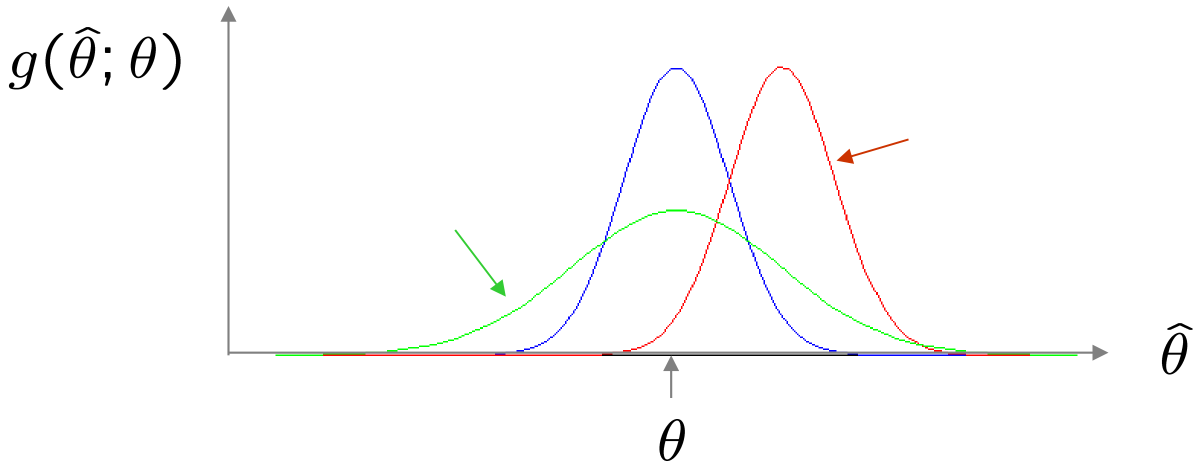
$\widehat{\mu} = 42$

$\widehat{\mu} = (\min(x_i) + \max(x_i))/2$ Mittelwert des größten und kleinsten Wertes

$\widehat{\mu} = \text{Median der Stichprobe}$

Eigenschaften eines Schätzers

Aus Wiederholung von Messungen ergibt sich WDF für:



Wir wollen kleinen (oder Null) (syst. Fehler): $b = E[\widehat{\theta}] - \theta$

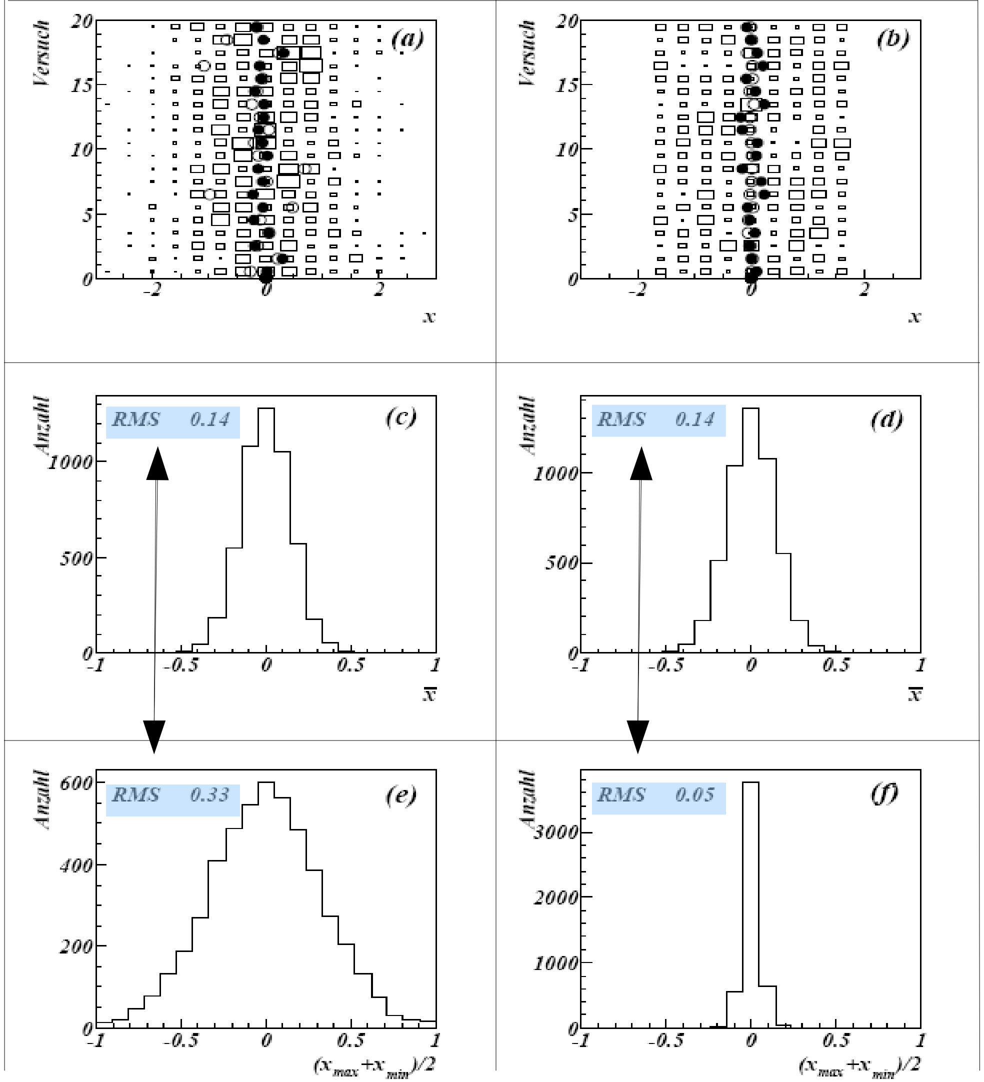
→ Mittelwert aus wiederholten Messungen gegen wahren Wert.

Und wir wollen kleine Variance (statistische Fehler): $V[\widehat{\theta}]$

→ kleiner Bias & Varianz konkurrierende Kriterien

Vergleich der Effizienz für Schätzer des Mittelwertes mean

	Gauss WDF	Gleichverteilt WDF (-2 < x_i < 2)



x_i

$\overline{x}$

$\overline{x'} = (x_{max} - x_{min})/2$

$\overline{x}$ deutlich effizienter als $\overline{x'}$ für Gauss PDF

$\overline{x'}$ deutlich Effizienter als $\overline{x}$ für Gleichv. (!)

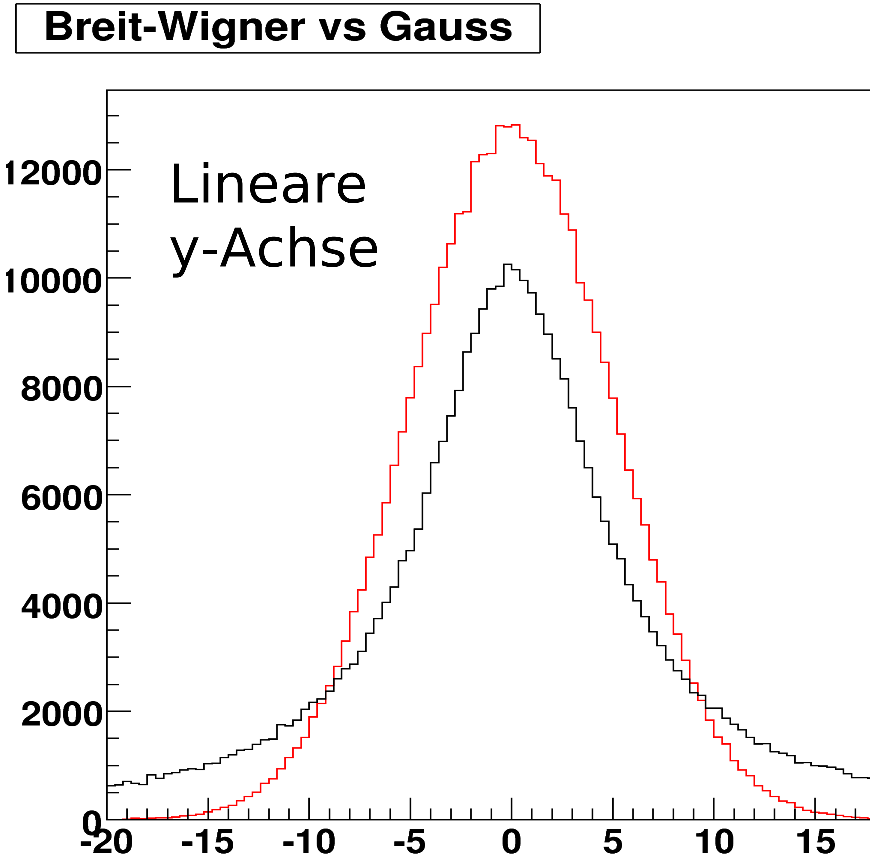
Robustheit: „truncated mean“

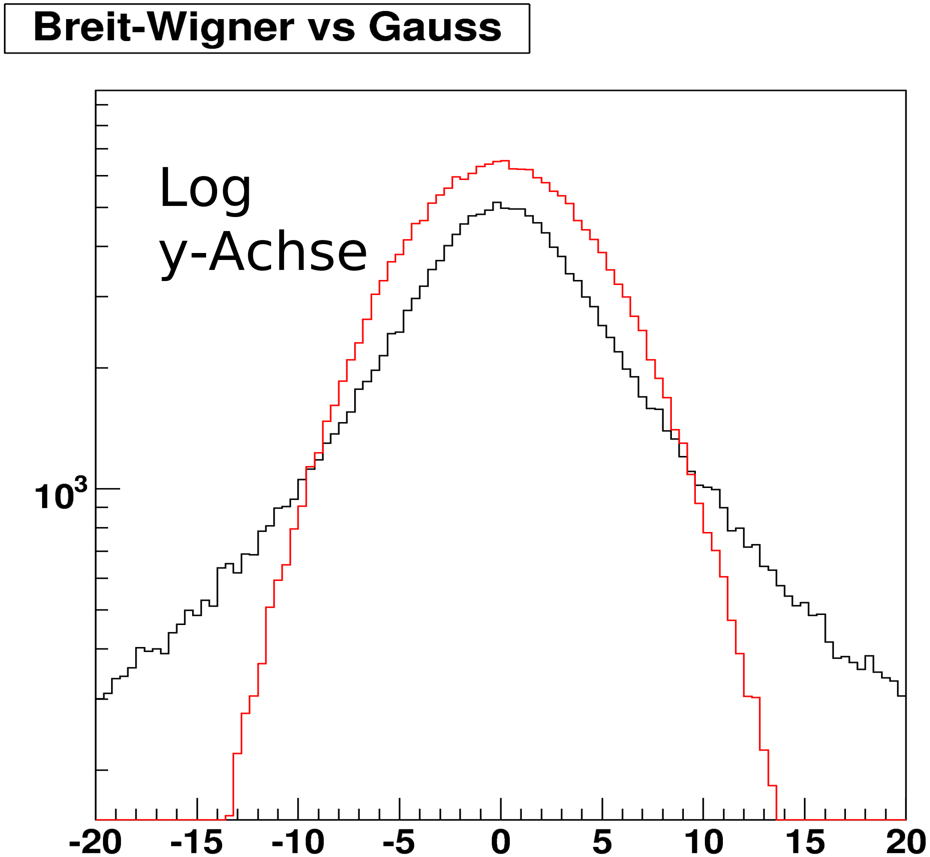
Gaussian pdf Breit-Wigner pdf

Ungefähr gleiche Varianz in Wertebereich

Aufgabe: gute Schätzer finden für die Symmetriepunkt von der Verteilung (x=0)

Aritmetische Mittelwert?

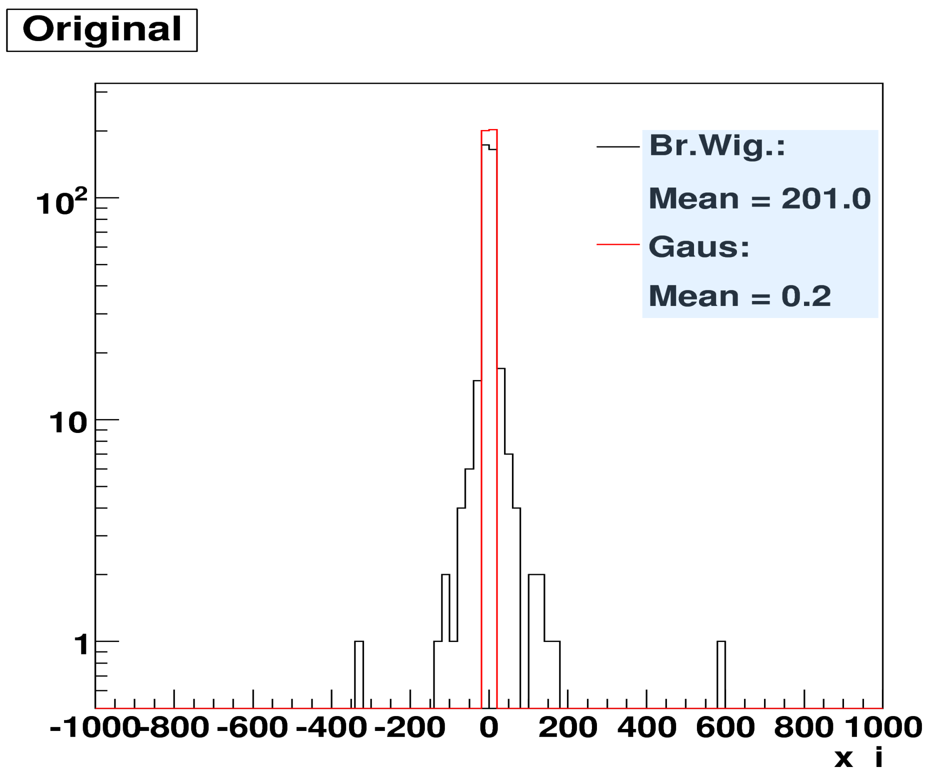




Robustheit: „truncated mean“

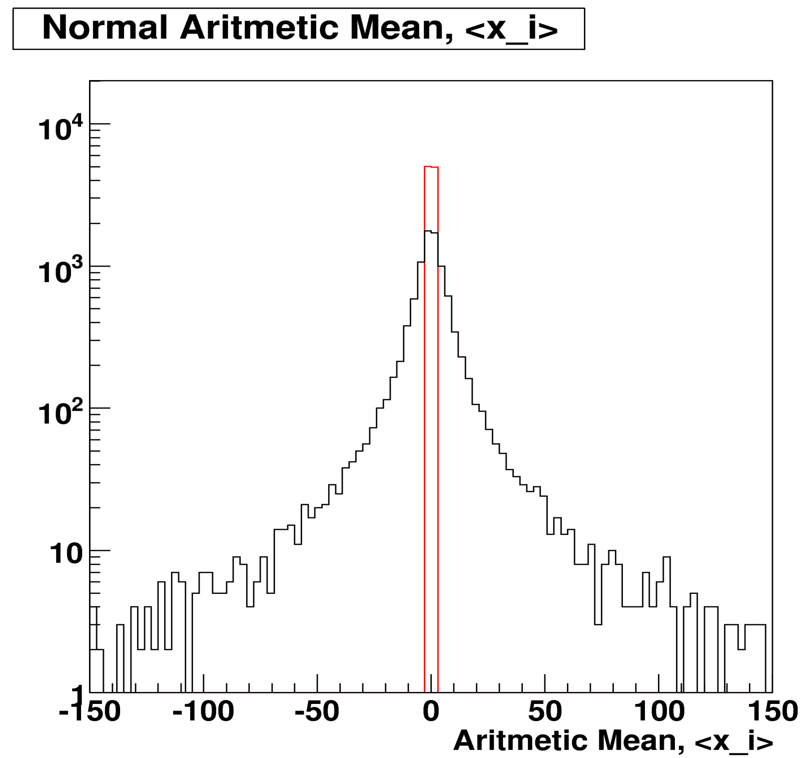
Stichprobe von 101 Zufallszahlen x_i von jeder PDF.

Breit Wigner: sehr grosse Werte |x_i| führt zu grosse
Werte für aritmetische Mittlewert <x_i>.



100 000 Stichproben
==> WDF für aritmetische Mittelwert, <x_i>

Breit Wigner: viel grössere Varianz als für Gaus
wegen Stichproben mit grosse Einzelwerte x_i
(siehe unten, links)

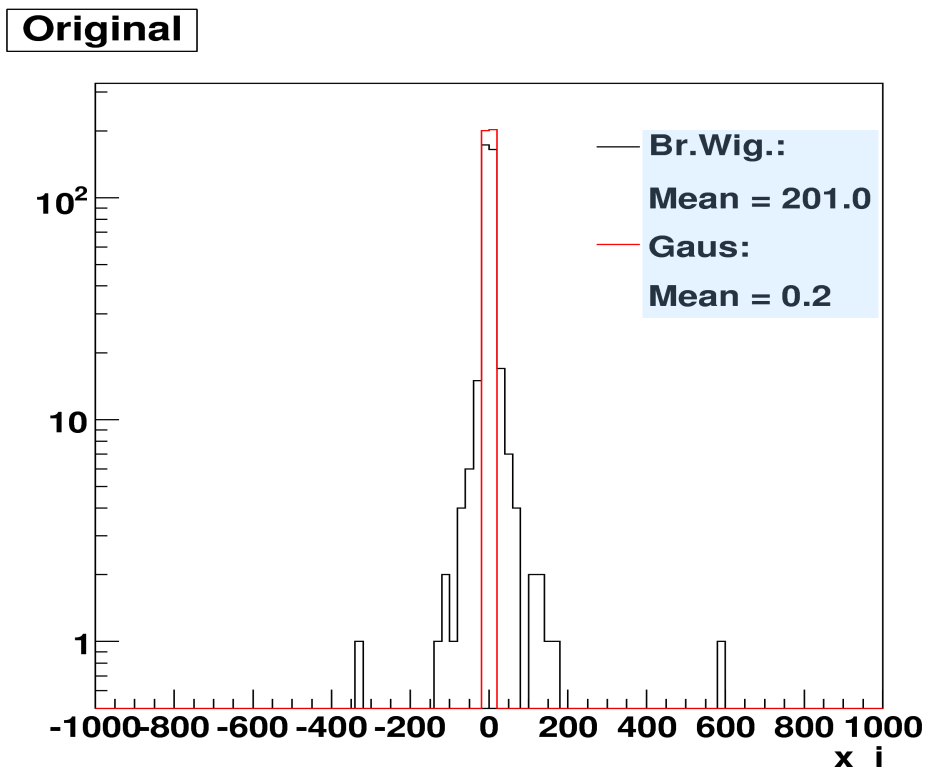


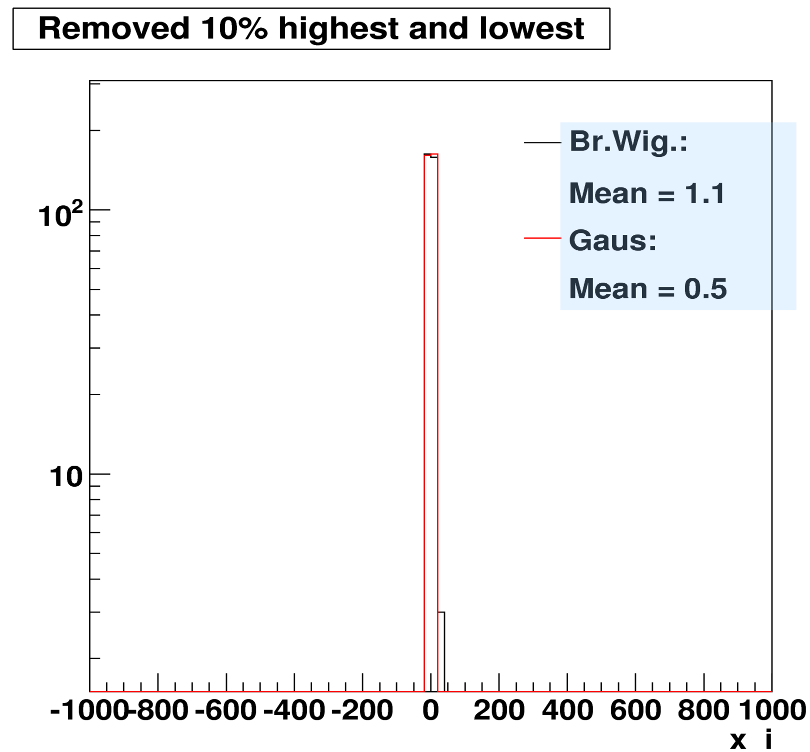
Stichprobe von 101 Zufallszahlen x_i von jeder PDF.

Breit Wigner: sehr grosse Werte |x_i| führt zu grosse Werte für aritmetische Mittlewert <x_i>.

Verbesserte Schätzer für Symmetriepunkt für Breit Wigner: 10% grösste und kleinste Werte für x_i ignorieren, dann <x_i> berechnen (Englisch: “truncated mean”)

==> viel kleinere Varianz für <x_i> für Breit Wigner

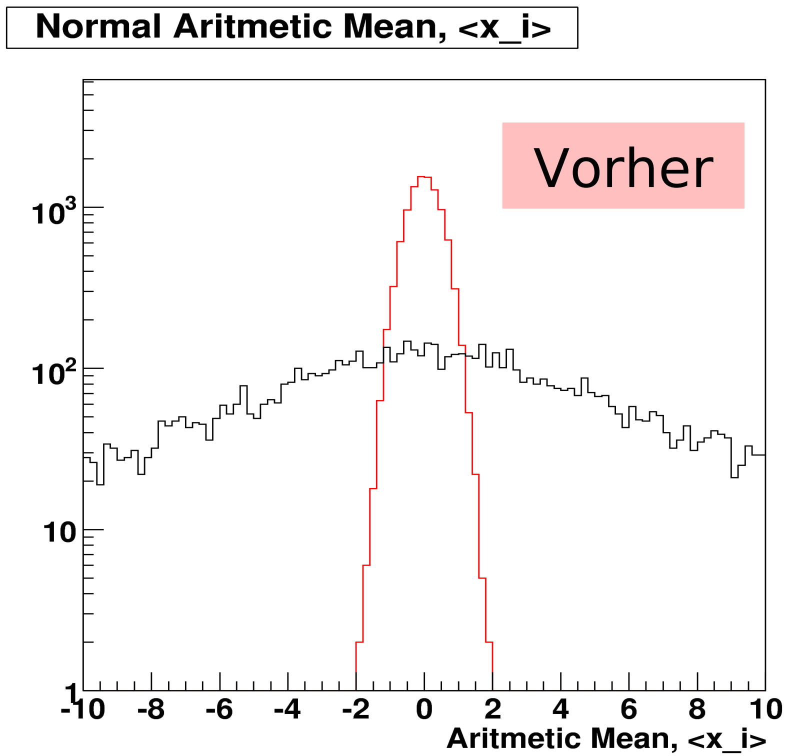




Robustheit: „truncated mean“

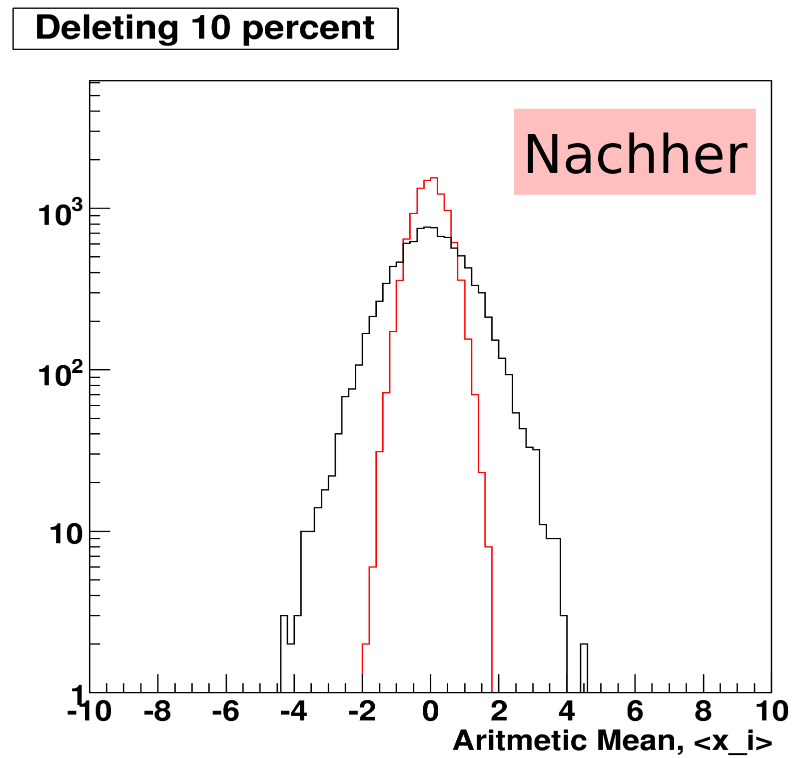
WDF für aritmetische Mittelwert, <x_i>

NORMAL

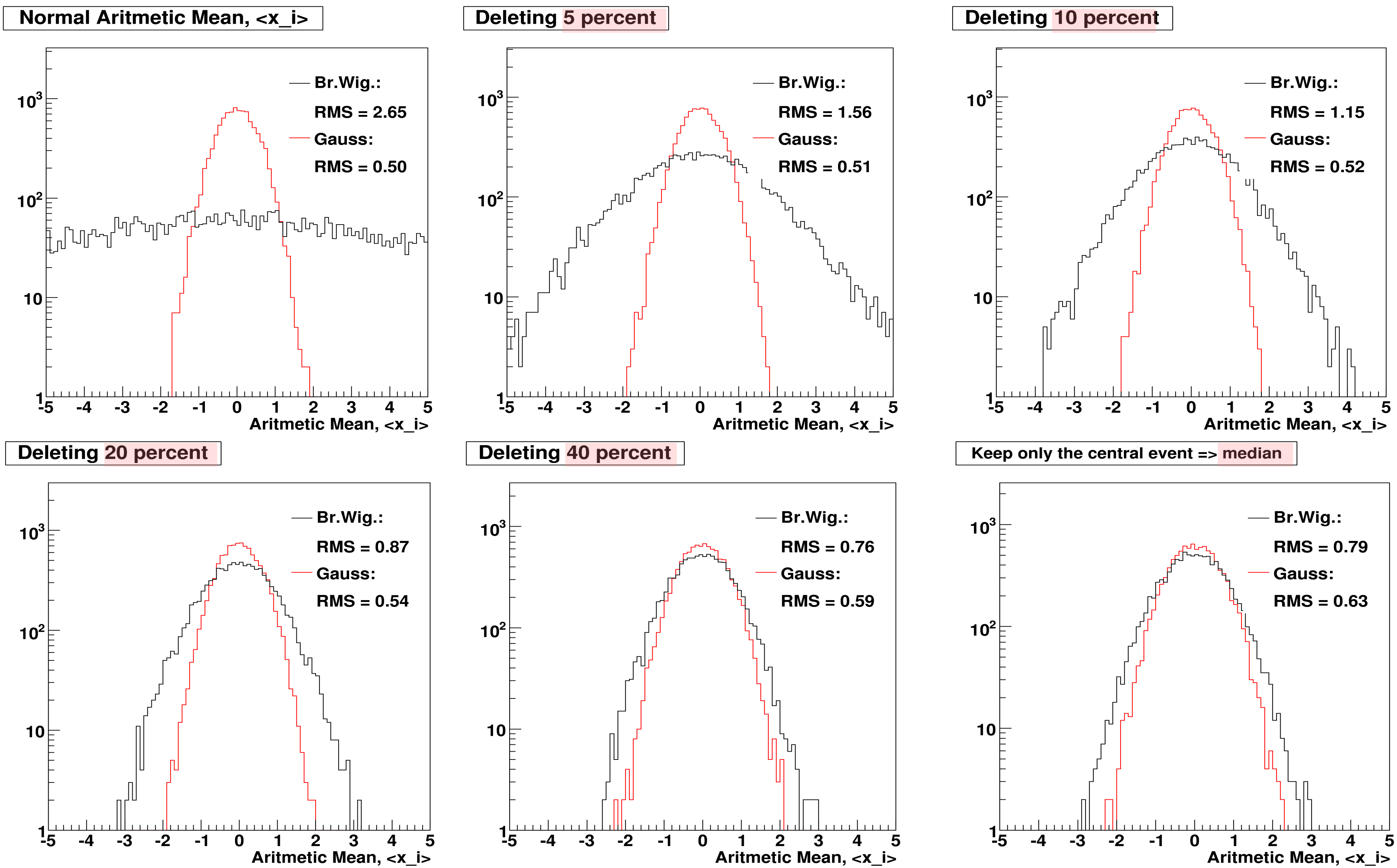


WDF für aritmetische Mittelwert, <x_i>

10% grösste und kleinste Werte ignoriert



Robustheit: „truncated mean“



Robustheit: „truncated mean“

RMS von <x_i> als Funktion von %-Anteil ignoriert