

# Experimentelle Methoden der Teilchenphysik

Sommersemester 2011/2012

Albert-Ludwigs-Universität Freiburg



Prof. Markus Schumacher

Physikalisches Institut, Westbau, 2. OG Raum 008

Telefon 07621 203 7612

E-Mail: [Markus.Schumacher@physik.uni-freiburg.de](mailto:Markus.Schumacher@physik.uni-freiburg.de)

Kapitel 14: Vertrauensintervalle

<http://terascale.physik.uni-freiburg.de/lehre/Sommersemester%202012>

# Schätzung von Vertrauensintervallen

Bisher: Schätzung von Parametern und deren Varianz

Jetzt: Angabe eines Intervalls “Vertrauensintervall”  $[a,b]$ , welches irgendeine Wahrscheinlichkeitsaussage über Beziehung zwischen statistischer Genauigkeit der Messung und dem wahren Wert des Parameters macht.

oft wird  $\pm 1$  Standardabweichung als Vertrauensintervall(Konfidenzintervall zu 68%Vertrauensniveau/Konfidenzniveau (CL)angegeben

aber: - nimmt an dass die WDF des Schätzers eine Gaussverteilung ist  
- Schwierigkeit bei Schätzwerten nahe an Physikalischen Grenzen  
z.B. - Masse =  $-5 \pm 2 < 0$  - Zählrate nahe bei oder  $< 0$

Deutliche Unterscheidung in Interpretation des Intervalls  $[a,b]$   
in den beiden Statistikschulen: Frequentisten (klassisch)  
Bayesianer (subjektiv).

# Interpretation von Vertrauensintervallen

Vertrauensintervall hängt von Messwerten ab  $\rightarrow [a,b]$  ist Zufallsvariable

Frequentisten: Vertrauensintervall  $[a,b]$  zu  $xy\%$  Vertrauensniveau CL

Interpretation: bei sehr häufiger Wiederholung identischer Messungen und Bestimmung der Vertrauensintervalle nach selber Methode, wird der Anteil von  $xy\%$  der so gebildeten Vertrauensintervalle den wahren Wert des Parameters enthalten (Abdeckungswkt.)

- keine Wahrscheinlichkeitsaussage über Beziehung von wahren Wert und Intervall in einzelner Messung
- kein Problem mit Intervallen im unphysikalischen Bereich.
- Abdeckungswahrscheinlichkeit = CL

Bayesianer: Glaubwürdigkeitsintervall  $[a,b]$  zu  $xy\%$  CL

Interpretation: die Wahrscheinlichkeit "Grad an Glaube", dass der wahre Wert im geschätzten Intervall liegt ist  $xy\%$

- Problem mit Intervallen im unphysikalischen Bereich
- Verlangt A-Priori-Wkt. für wahre Werte des Parameters

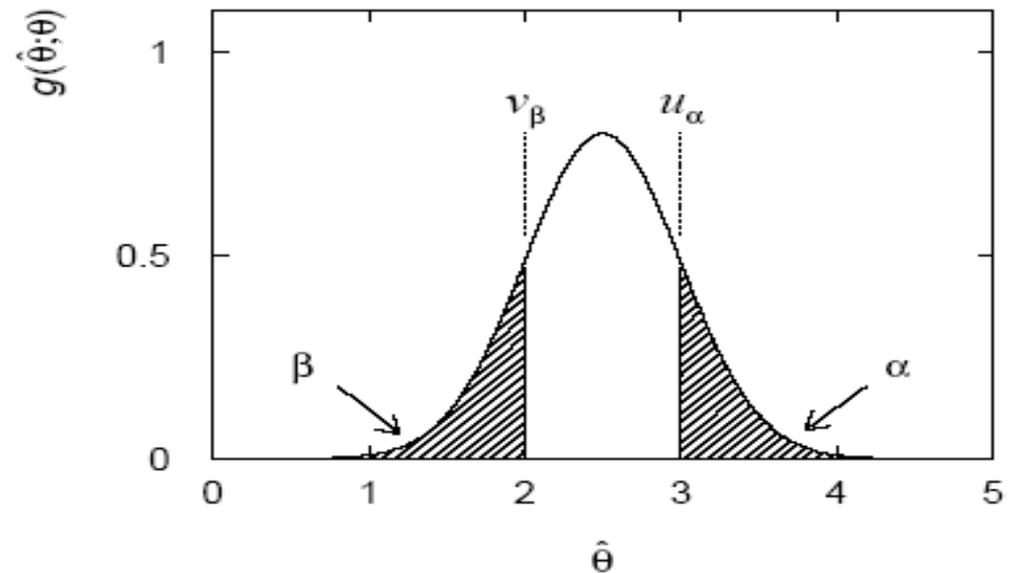
# Bestimmung Frequentistischer Vertrauensintervalle

Betrachte einen Schätzer  $\hat{\theta}$  für einen Parameter  $\theta$  und einen Schätzwert, der aus einer Stichprobe gewonnen wird  $\hat{\theta}_{\text{obs}}$ .

Man braucht für alle möglichen wahren Parameter  $\theta$  die Wahrscheinlichkeitsdichtefunktion für den Schätzer  $g(\hat{\theta}; \theta)$ .

Spezifiziere obere und untere Ausläuferwahrscheinlichkeiten z.B.  $\alpha = 0.05$ ,  $\beta = 0.05$ , und bestimme die Funktionen  $u_\alpha(\theta)$  and  $v_\beta(\theta)$  so das gilt:

$$\begin{aligned}\alpha &= P(\hat{\theta} \geq u_\alpha(\theta)) \\ &= \int_{u_\alpha(\theta)}^{\infty} g(\hat{\theta}; \theta) d\hat{\theta} \\ \beta &= P(\hat{\theta} \leq v_\beta(\theta)) \\ &= \int_{-\infty}^{v_\beta(\theta)} g(\hat{\theta}; \theta) d\hat{\theta}\end{aligned}$$



# Bestimmung Frequentistischer Vertrauensintervalle

Die Region zwischen  $u_\alpha(\theta)$  und  $v_\beta(\theta)$  wird der Vertrauensgürtel genannt

$$P(l_\beta(\theta) \leq \hat{\theta} \leq u_\alpha(\theta)) = 1 - \alpha - \beta$$

Finde die Punkte, wo der beobachtete Schätzwert den Konfidenzgürtel schneidet  $\rightarrow [a, b]$

$$a(\hat{\theta}) \equiv u_\alpha^{-1}(\hat{\theta})$$

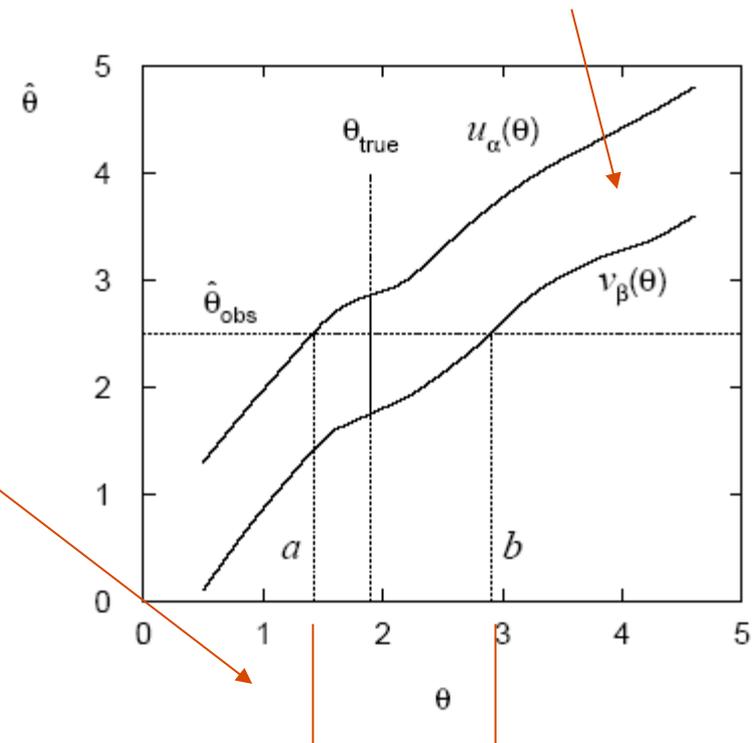
$$b(\hat{\theta}) \equiv l_\beta^{-1}(\hat{\theta}).$$

Es gilt für alle Werte  $\theta$ :

$$\hat{\theta} \geq u_\alpha(\theta) \Leftrightarrow a(\hat{\theta}) \geq \theta \quad P(a(\hat{\theta}) \geq \theta) = \alpha$$

$$\hat{\theta} \leq l_\beta(\theta) \Leftrightarrow b(\hat{\theta}) \leq \theta, \quad P(b(\hat{\theta}) \leq \theta) = \beta$$

$$P(a(\hat{\theta}) \leq \theta \leq b(\hat{\theta})) = 1 - \alpha - \beta.$$



Konfidenzniveau =  $1 - \alpha - \beta$  = Abdeckwahrscheinlichkeit des Intervalls den wahren Wert zu enthalten (in frequentistischer Interpretation s.o.)

Gilt für jeden beliebigen wahren Wert des Parameters

# KI aus Invertierung eines Hypothesentests

Das Konfidenzintervall für einen Parameter  $\theta$  kann durch Definition eines Tests für den hypothetischen Wert  $\theta$  gefunden werden (und mache dies für alle hypothetischen Werte  $\theta$ ):

Spezifiziere Werte der Daten die unverträglich mit  $\theta$  sind (Kritische Region) so dass die Wahrscheinlichkeit  $P(\text{Daten in kritischer Region}) \leq \gamma$  für eine vorgegebenes  $\gamma$ , z.B., 0.05 or 0.1.

Wenn beobachtete Daten in kritischer Region, dann verwerfe den Wert von  $\theta$ .

Nun invertiere den Test, um eine Konfidenzintervall zu definieren gemäß:

Menge der  $\theta$ -Werte, die nicht verworfen werden würden in einem Hypothesentest mit Signifikanzniveau  $\gamma$  (Konfidenzniveau ist  $1 - \gamma$ ).

Das Intervall wird den wahren Wert  $\theta$  mit Wkt  $\geq 1 - \gamma$  enthalten. Dies ist Äquivalent zur Konstruktion des Konfidenzgürtels: der Konfidenzgürtel ist die Akzeptanzregion eines Hypothesentests.

# Beziehung zwischen CL und P-Wert

Äquivalente Sichtweise: wir betrachten einen Signifikanztest für jeden hypothetischen Wert von  $\theta$ , mit dem Ergebnis des  $p$ -Wertes,  $p_\theta$ .

Wenn  $p_\theta < \gamma$ , dann verwerfe  $\theta$ .

Das Konfidenzintervall zum Konfidenzniveau  $CL = 1 - \gamma$  beinhaltet die Werte für  $\theta$ , die nicht verworfen werden.

Z.B. eine untere Grenze/Limit  $a$  auf  $\theta$  ist der größte Wert für  $\theta$  für den gilt  $p_\theta \geq \gamma$ .

In der Praxis finden wir den Grenzwert indem wir  $p_\theta = \gamma$  setzen und diese Beziehung für  $\theta$  lösen.

Oder für Signifikanzniveaus  $\alpha$  u.  $\beta$  größtes  $a$  und kleinstes  $b$  finden, dass folgende Gleichungen erfüllt

$$\alpha = \int_{\hat{\theta}_{obs}}^{\infty} g(\hat{\theta}; a) d\hat{\theta} = 1 - G(\hat{\theta}_{obs}; a)$$

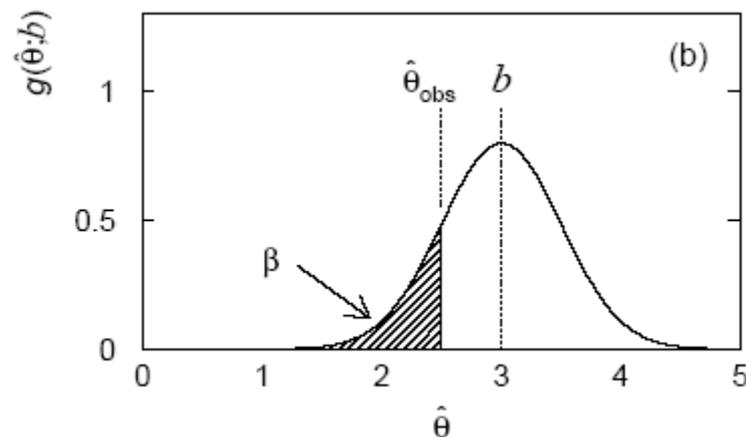
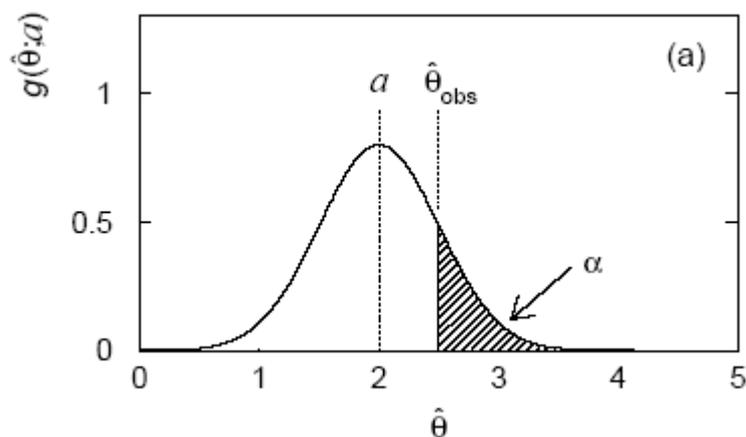
$$\beta = \int_{-\infty}^{\hat{\theta}_{obs}} g(\hat{\theta}; b) d\hat{\theta} = G(\hat{\theta}_{obs}; b).$$

# Konfidenzintervalle in der Praxis

Das Rezept das Intervall  $[a, b]$  zu finden, reduziert sich darauf zu lösen

$$\alpha = \int_{u_\alpha(\theta)}^{\infty} g(\hat{\theta}; \theta) d\hat{\theta} = \int_{\hat{\theta}_{\text{obs}}}^{\infty} g(\hat{\theta}; a) d\hat{\theta},$$

$$\beta = \int_{-\infty}^{v_\beta(\theta)} g(\hat{\theta}; \theta) d\hat{\theta} = \int_{-\infty}^{\hat{\theta}_{\text{obs}}} g(\hat{\theta}; b) d\hat{\theta}.$$

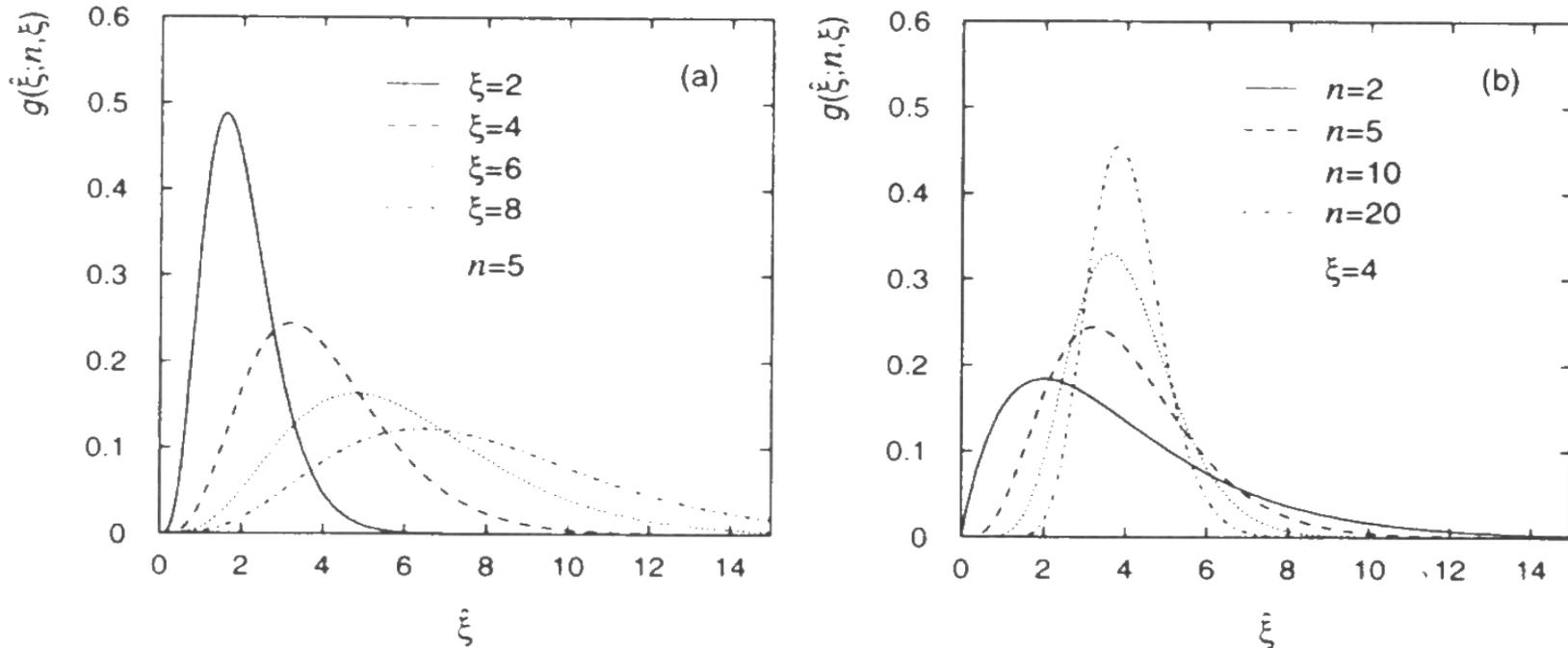


→  $a$  ist hypothetischer Wert von  $\theta$  so dass  $P(\hat{\theta} > \hat{\theta}_{\text{obs}}) = \alpha$ .

→  $b$  ist hypothetischer Wert von  $\theta$  so dass  $P(\hat{\theta} < \hat{\theta}_{\text{obs}}) = \beta$ .

# Beispiel Schätzer für Exponential-WDF

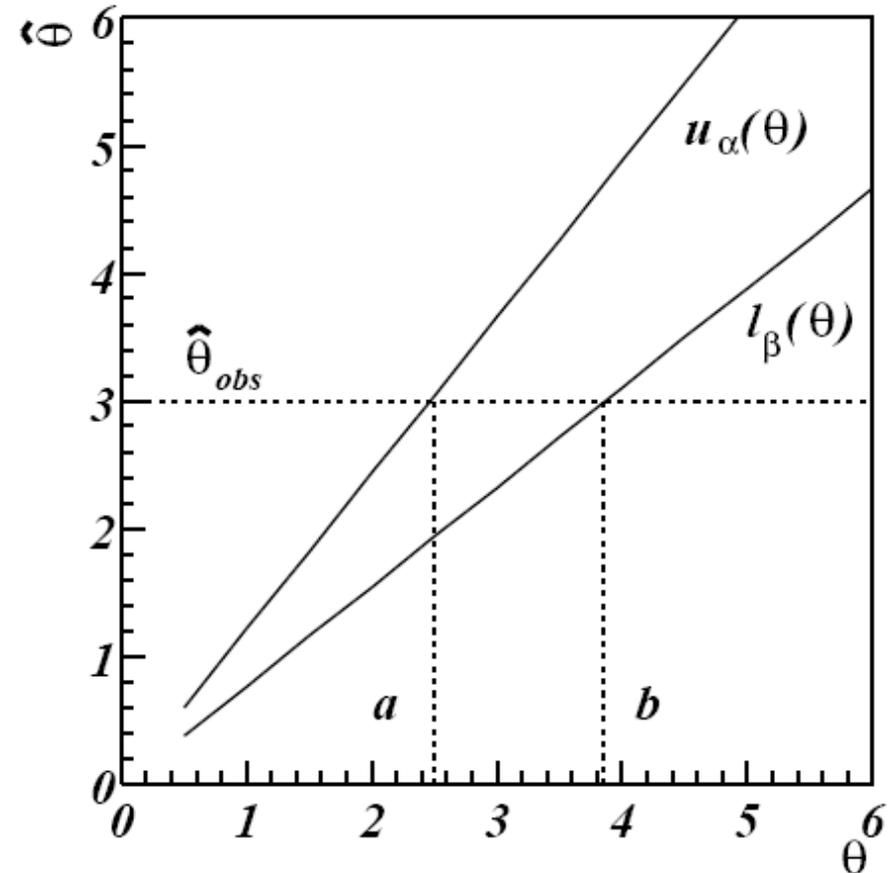
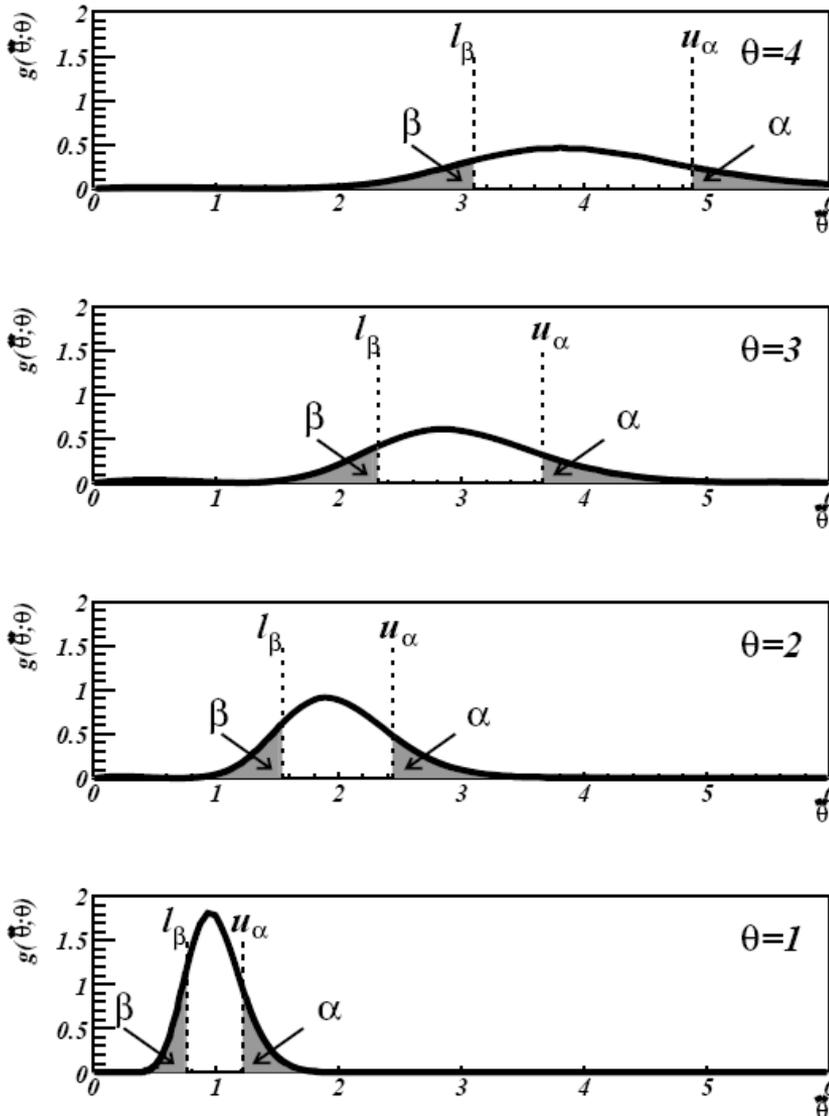
WDF für den ML-Schätzer = arithmetischer Mittelwert der Lebensdauern



**Fig. 10.1** The sampling p.d.f.  $g(\hat{\xi}; n, \xi)$  for the estimator  $\hat{\xi}$  for various values of  $n$  and  $\xi$ . (a)  $n = 5$  measurements and various values of the true parameter  $\xi$ . (b)  $\xi = 4$  and various numbers of measurements  $n$ .

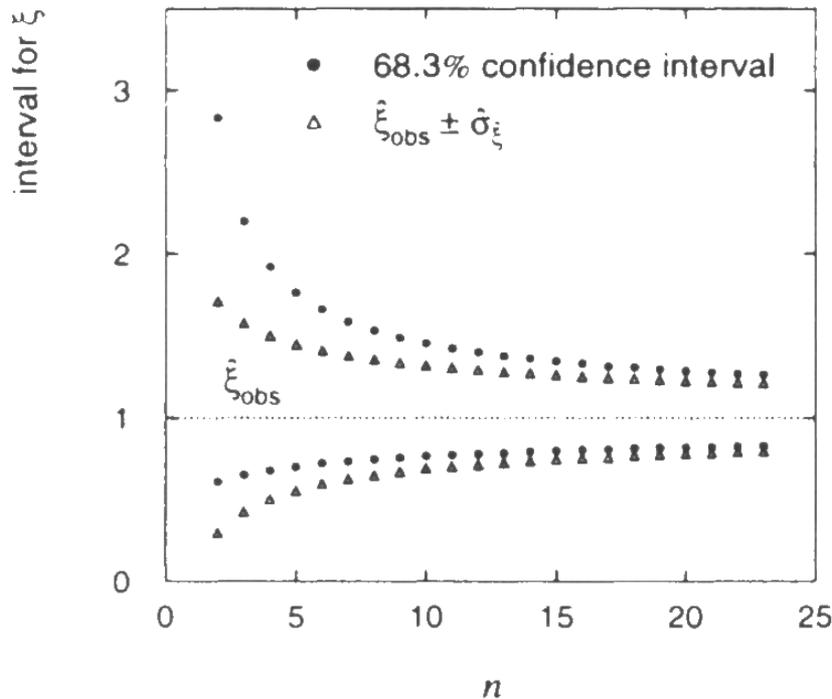
# Beispiel Schätzer für Exponential-WDF

$n = 20$  Messungen



# Beispiel Schätzer für Exponential-WDF

Vergleich von KI aus beobachteten Schätzwert  $\pm 1$  Standardabweichung und korrekter Berechnung aus Konfidenzgürtel



**Fig. 10.2** Classical confidence intervals for the parameter of the exponential distribution  $\xi$  (between solid points) and the interval  $[\hat{\xi}_{obs} - \hat{\sigma}_{\hat{\xi}}, \hat{\xi}_{obs} + \hat{\sigma}_{\hat{\xi}}]$  (between open triangles) for different values of the number of measurements  $n$ , assuming an observed value  $\hat{\xi}_{obs} = 1$ .

asymptotisch identisch, da WDF für Schätzer gegen Gauss-WDF  
für kleine Stichproben: korrekte KI asymmetrisch u. größer  
KI aus  $\pm 1$  Standardabweichung zu kleine Abdeckungswahrscheinlichkeit

# Zentrale und einseitige Konfidenzintervalle

Manchmal wird nur  $\alpha$  oder  $\beta$  angegeben  $\rightarrow$  einseitiges Intervall  
(Ausschlussgrenze oder Limit genannt)

Freiheit: wie  $\gamma$  auf  $\alpha$  und  $\beta$  in den rechten u. linken Ausläufern verteilen?

Meistens wird  $\alpha=\beta=\gamma/2$  genommen  $\rightarrow$  Abdeckungswkt.  $1-\gamma$   
 $\rightarrow$  zentrales Konfidenzintervall

aber zentrales Intervall bedeutet nicht, dass es symmetrisch um Schätzwert ist, i.e.  $b-\theta \neq \theta -a$

Es gibt auch: symmetrische KI:  $b - \theta = \theta -a$   
kürzeste KI: minimale Länge von  $[a,b]$

Meist Konvention:  $CL = 1-\gamma = 68,3\%$

# KI für Gaussverteilte Schätzer

$$g(\hat{\theta}; \theta) = \frac{1}{\sqrt{2\pi\sigma_{\hat{\theta}}^2}} \exp\left(-\frac{(\hat{\theta} - \theta)^2}{2\sigma_{\hat{\theta}}^2}\right)$$

oft im Grenzfall erfüllt  
(großen Stichprobenumfang, zentraler Grenzwertsatz)  
Annahme: Varianz unabhängig von Parameter

Ergebnis für CL=68.3%:  $\theta = \hat{\theta}_{obs} \pm \hat{\sigma}_{\hat{\theta}}$

Formale Berechnung:

Kumulativfunktion der WDF für den Schätzer::

$$G(\hat{\theta}; \theta, \sigma_{\hat{\theta}}) = \int_{-\infty}^{\hat{\theta}} \frac{1}{\sqrt{2\pi\sigma_{\hat{\theta}}^2}} \exp\left(-\frac{(\hat{\theta}' - \theta)^2}{2\sigma_{\hat{\theta}}^2}\right) d\hat{\theta}' = \Phi\left(\frac{\hat{\theta} - \theta}{\sigma_{\hat{\theta}}}\right)$$

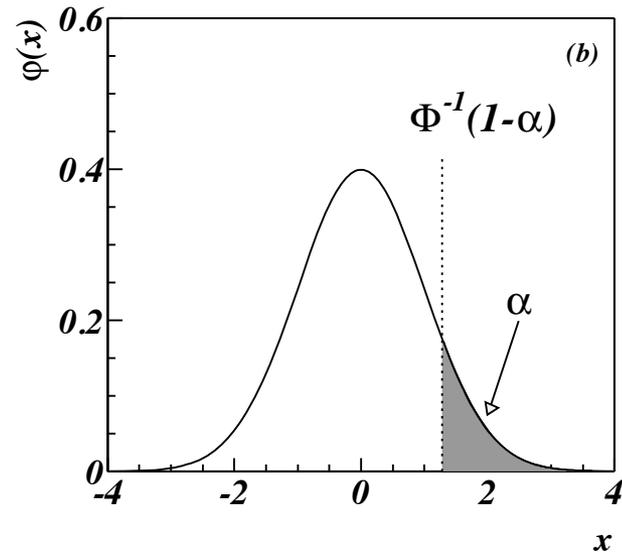
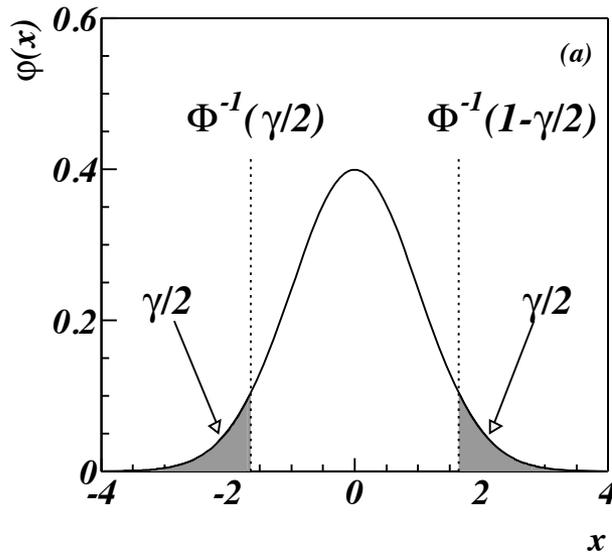
Bestimmungsgleichungen für a und b:

$$\alpha = 1 - G(\hat{\theta}_{obs}; a, \sigma_{\hat{\theta}}) = 1 - \Phi\left(\frac{\hat{\theta}_{obs} - a}{\sigma_{\hat{\theta}}}\right)$$
$$\beta = G(\hat{\theta}_{obs}; b, \sigma_{\hat{\theta}}) = \Phi\left(\frac{\hat{\theta}_{obs} - b}{\sigma_{\hat{\theta}}}\right),$$

und deren Lösung für a und b:

$$a = \hat{\theta}_{obs} - \sigma_{\hat{\theta}}\Phi^{-1}(1 - \alpha)$$
$$b = \hat{\theta}_{obs} + \sigma_{\hat{\theta}}\Phi^{-1}(1 - \beta).$$

# KI für Gaussverteilte Schätzer



Konfidenzniveaus für Intervalle			
zentrale		einseitige	
$\Phi^{-1}(1 - \gamma/2)$	$1 - \gamma$	$\Phi^{-1}(1 - \alpha)$	$1 - \alpha$
1	0.6827	1	0.8413
2	0.9544	2	0.9772
3	0.9973	3	0.9987
4	$1 - 6.3 \times 10^{-5}$		
5	$1 - 5.7 \times 10^{-7}$		
6	$1 - 2.0 \times 10^{-9}$		

Quantile für Intervalle			
zentrale		einseitige	
$1 - \gamma$	$\Phi^{-1}(1 - \gamma/2)$	$1 - \alpha$	$\Phi^{-1}(1 - \alpha)$
0.90	1.645	0.90	1.282
0.95	1.960	0.95	1.645
0.99	2.576	0.99	2.326
0.999	3.29		
0.9999	3.89		

# KI aus der Likelihoodfunktion

Im Grenzfall großen Stichprobenumfangs gilt,  
dass die WDF für die ML-Schätzer einer n-dimensionalen Gauss-WDF mit  
Varianz V folgt:

$$g(\hat{\vec{\theta}}; \vec{\theta}) = \frac{1}{(2\pi)^{n/2} |V|^{1/2}} \exp\left(-\frac{1}{2} Q(\hat{\vec{\theta}}, \vec{\theta})\right) \quad Q(\hat{\vec{\theta}}, \vec{\theta}) = (\hat{\vec{\theta}} - \vec{\theta})^T V^{-1} (\hat{\vec{\theta}} - \vec{\theta}).$$

Die Likelihoodfunktion in der Nähe des Maximums wird zur n-dim.  
Gaussfunktion, die log-Likelihood-Fkt zum ein Hyperparaboloid:

$$L(\vec{\theta}) = L_{\max} \exp\left[-\frac{1}{2} Q(\hat{\vec{\theta}}, \vec{\theta})\right], \quad Q(\hat{\vec{\theta}}, \vec{\theta}) = (\hat{\vec{\theta}} - \vec{\theta})^T V^{-1} (\hat{\vec{\theta}} - \vec{\theta})$$

$Q(\hat{\vec{\theta}}, \vec{\theta}) = Q_\gamma$  definiert eine hyper-ellipsoidale Konfidenzregion

$$P(\text{ellipsoid covers true } \vec{\theta}) = P(Q(\hat{\vec{\theta}}, \vec{\theta}) \leq Q_\gamma)$$

Wenn  $\hat{\vec{\theta}} \sim N(\vec{\theta}, V)$  dann  $Q(\hat{\vec{\theta}}, \vec{\theta}) \sim \text{Chi-square}(n)$

$$\text{coverage probability} \equiv 1 - \gamma = \int_0^{Q_\gamma} f_{\chi^2}(z; n) dz = F_{\chi^2}(Q_\gamma; n)$$

# Approximative KI aus $L(\theta)$

Das Rezept die Konfidenzregion zum Konfidenz  $CL = 1-\gamma$  zu finden, lautet

$$\ln L(\vec{\theta}) = \ln L_{\max} - \frac{Q_\gamma}{2} \quad \text{or} \quad \chi^2(\vec{\theta}) = \chi_{\min}^2 + Q_\gamma$$

$$\text{where} \quad Q_\gamma = F_{\chi^2}^{-1}(1 - \gamma; n)$$

$Q_\gamma$	$1 - \gamma$				
	$n = 1$	$n = 2$	$n = 3$	$n = 4$	$n = 5$
1.0	0.683	0.393	0.199	0.090	0.037
2.0	0.843	0.632	0.428	0.264	0.151
4.0	0.954	0.865	0.739	0.594	0.451
9.0	0.997	0.989	0.971	0.939	0.891

$1 - \gamma$	$Q_\gamma$				
	$n = 1$	$n = 2$	$n = 3$	$n = 4$	$n = 5$
0.683	1.00	2.30	3.53	4.72	5.89
0.90	2.71	4.61	6.25	7.78	9.24
0.95	3.84	5.99	7.82	9.49	11.1
0.99	6.63	9.21	11.3	13.3	15.1

Für endliche Stichprobe sind dies approximative KI.

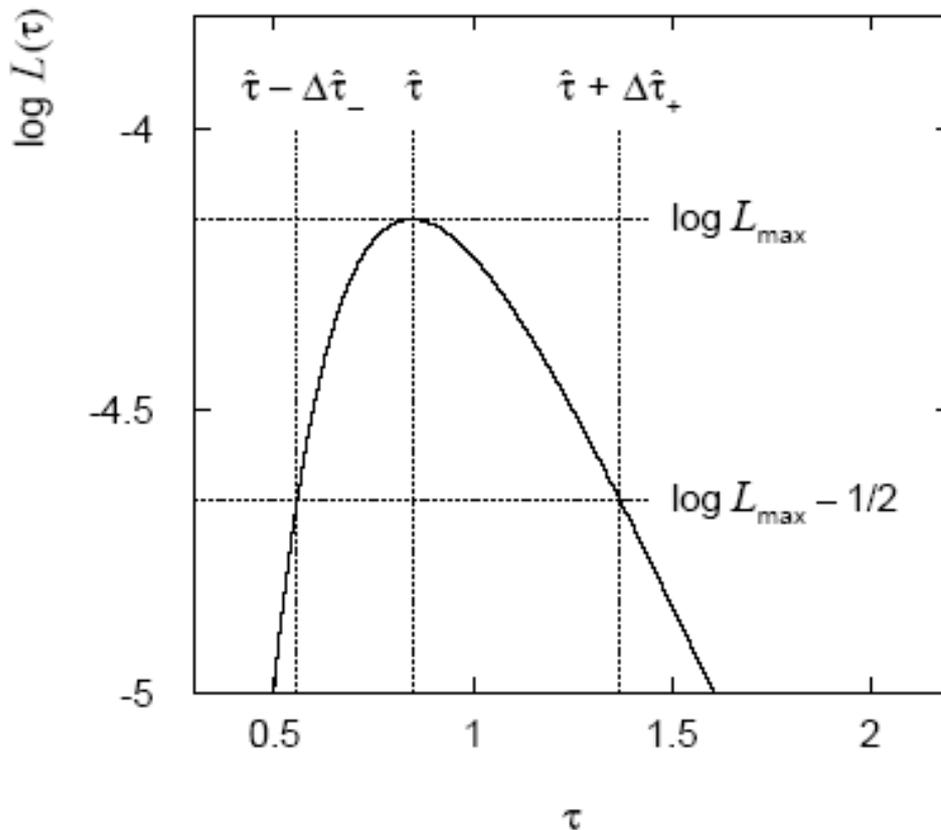
Nicht garantiert, dass Abdeckungswahrscheinlichkeit exakt  $1-\gamma$  ist;

es gibt kein einfaches Theorem welches die Qualität der Näherung beschreibt  
→ einziger Ausweg ist die MC-Simulation.

# Beispiel für approx. KI aus $\ln L(\theta)$

Für  $n=1$  Parameter,  $CL = 0.683$ ,  $Q_\gamma = 1$ .

Our exponential example, now with  $n = 5$  observations:

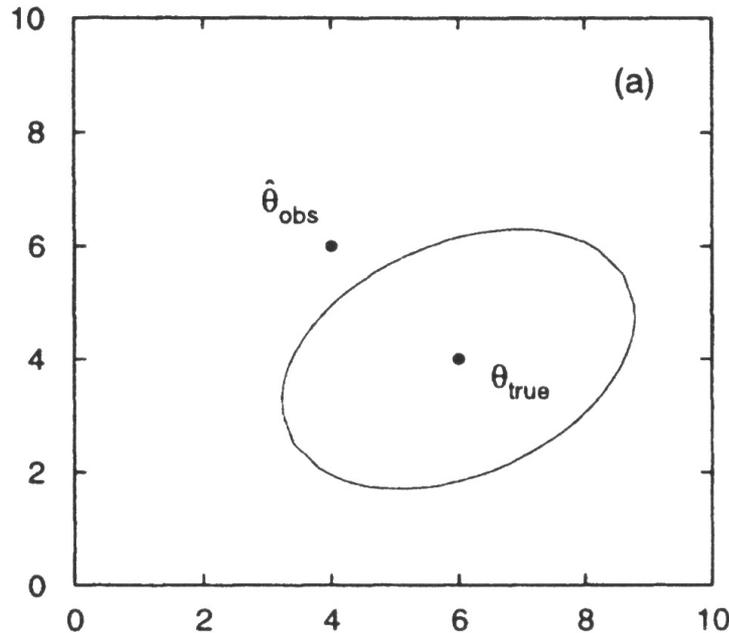


$$\hat{\tau} = 0.85^{+0.52}_{-0.30}$$

# 2-Dim Beispiel für KI aus $\ln L(\theta)$

$\Delta \ln L = 1/2$ -Kontur um wahren Wert

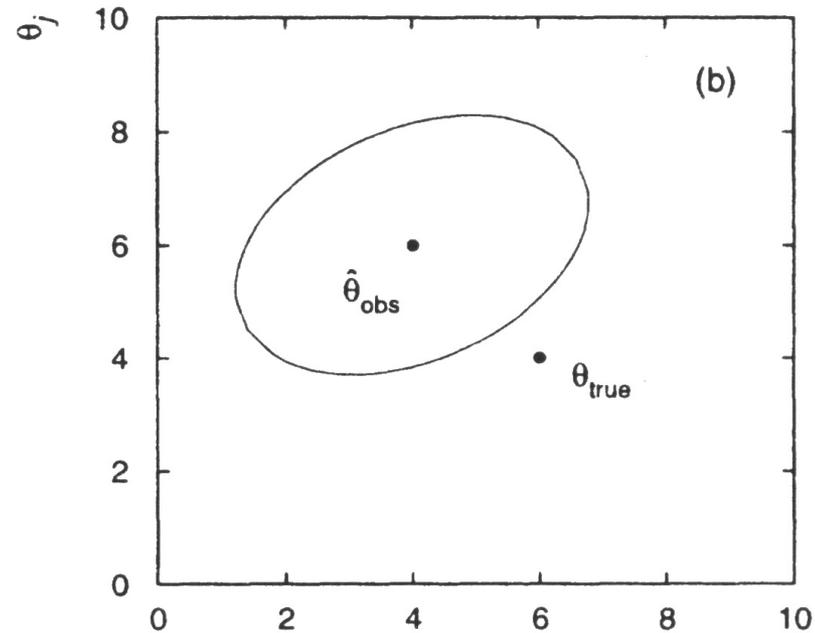
$$g(\hat{\theta}; \vec{\theta}) = \frac{1}{(2\pi)^{n/2} |V|^{1/2}} \exp\left(-\frac{1}{2} Q(\hat{\theta}, \vec{\theta})\right)$$



$\Delta \ln L = 1/2$ -Kontur um Schätzwert

= 2-dim Konfidenzintervall zu CL=39%

$$\mathcal{L}(\vec{\theta}) = \mathcal{L}_{\max} \exp\left(-\frac{1}{2} Q(\vec{\theta}, \hat{\theta})\right)$$



$$Q(\vec{\theta}_1, \vec{\theta}_2) \equiv (\vec{\theta}_1 - \vec{\theta}_2)^T V^{-1} (\vec{\theta}_1 - \vec{\theta}_2)$$

# KI für Poisson-Parameter

$$f(n; \lambda) = \frac{\lambda^n}{n!} \exp(-\lambda)$$

$n$  = Anzahl beobachtete Ereignisse = ML-Schätzer  $\hat{\lambda}$   
Ziel: Konfidenzintervall für  $\lambda$

Auf Grund der diskreten Natur der ZV ist  $\alpha = P(\hat{\lambda} \geq u_\alpha(\lambda))$   
nicht für alle Schätzer  $\hat{\lambda}$  exakt zu erfüllen

Modifikation der Gleichungen für Bestimmung des Konfidenzgürtels:

$$\alpha \geq P(\hat{\lambda} \geq u_\alpha(\lambda)) \quad \text{minimale } u_\alpha(\lambda) \qquad \beta \geq P(\hat{\lambda} \leq l_\beta(\lambda)) \quad \text{maximalem } l_\beta(\lambda)$$

$$\text{Und damit: } P(l_\beta(\lambda) < \hat{\lambda} < u_\alpha(\lambda)) > 1 - \alpha - \beta \qquad P(a \leq \lambda \leq b) \geq 1 - \alpha - \beta$$

$$P(\lambda \geq a) \geq 1 - \alpha$$

$$P(\lambda \leq b) \geq 1 - \beta.$$

Lösen der Bedingungsgleichungen ergibt Intervallgrenzen  $a$  und  $b$

$$\alpha = \sum_{n=n_{obs}}^{\infty} f(n; a) = 1 - \sum_{n=0}^{n_{obs}-1} f(n; a) = 1 - \sum_{n=0}^{n_{obs}-1} \frac{a^n}{n!} e^{-a},$$

$$\beta = \sum_{n=0}^{n_{obs}} f(n; b) = \sum_{n=0}^{n_{obs}} \frac{b^n}{n!} e^{-b}.$$

# KI für Poisson-Parameter

Falls kein Ereignis beobachtet wird ergibt sich nur obere Grenze:

$$\beta = e^{-b} \implies b = -\log \beta \quad \text{für CL=95\% damit:} \quad b = -\log(0.05) = 2.996 \approx 3.$$

Für allgemeinen Fall nutze Beziehung zwischen Poisson-WDF u. Chi<sup>2</sup>-WDF

$$\begin{aligned} \sum_{n=0}^{n_{obs}} \frac{\lambda^n}{n!} e^{-\lambda} &= \int_{2\lambda}^{\infty} f_{\chi^2}(z; n_{dof} = 2(n_{obs} + 1)) dz \\ &= 1 - F_{\chi^2}(2\lambda; n_{dof} = 2(n_{obs} + 1)), \end{aligned}$$

Die Grenzen ergeben sich aus der kumulativen Chi-Quadrat-WDF zu:

$$\begin{aligned} a &= \frac{1}{2} F_{\chi^2}^{-1}(\alpha; n_{dof} = 2n_{obs}), \\ b &= \frac{1}{2} F_{\chi^2}^{-1}(1 - \beta; n_{dof} = 2(n_{obs} + 1)). \end{aligned}$$

$n_{obs}$	untere Schranke $a$		obere Schranke $b$	
	$\alpha = 0.1$ CL = 90%	$\alpha = 0.05$ CL = 95%	$\beta = 0.1$ CL = 90%	$\beta = 0.05$ CL = 95%
0	-	-	2.30	3.00
1	0.105	0.051	3.89	4.74
2	0.532	0.355	5.32	6.30
3	1.10	0.818	6.68	7.75
4	1.74	1.37	7.99	9.15
5	2.43	1.97	9.27	10.51
6	3.15	2.61	10.53	11.84
7	3.89	3.29	11.77	13.15
8	4.66	3.98	12.99	14.43
9	5.43	4.70	14.21	15.71
10	6.22	5.43	15.41	16.96

# Grenzen für Poisson-Parameter mit Untergrund

Betrachte wieder den Fall, dass wir  $n = n_s + n_b$  Ereignisse beobachten

$n_b$  Ereignisse aus bekannten Prozessen (Untergrund)

$n_s$  Ereignisse aus einem neuen Prozess (Signal)

sind Poisson-ZV mit Mittelwerten  $s, b$ , und daher folgt  $n = n_s + n_b$  ebenfalls einer Poisson-WDF mit Mittelwert  $= s + b$ . Annahme  $b$  bekannt.

Annahme: wir suchen nach Evidenz für den Signalprozess.  
aber die Anzahl der beobachteten Ereignisse ist in etwa gleich der Anzahl der Ereignisse, die wir von Untergrundereignissen erwarten  
z.B. Untergrunderwartung  $b=4.6$ ,  $n_{\text{obs}} = 5$  Ereignisse.

Der Hinweis für die Anwesenheit von Signalereignissen ist nicht statistisch signifikant,

→ setze einer obere Grenze auf den Parameter  $s$ .

# Grenzen für Poisson-Parameter mit Untergrund

Finde den hypothetischen Wert von  $s$ , so dass es eine kleine Wahrscheinlichkeit gibt, z.B.  $\gamma = 0.05$ , so wenige Ereignisse zu finden wie wir beobachten haben oder weniger:

$$\gamma = P(n \leq n_{\text{obs}}; s, b) = \sum_{n=0}^{n_{\text{obs}}} \frac{(s+b)^n}{n!} e^{-(s+b)}$$

Löse die Bedingung numerisch für  $s = s_{\text{up}}$ , dies ergibt die obere Ausschlussgrenze für  $s$  zu einem Vertrauensniveau  $1-\gamma$ .

Beispiel: nimm an  $b = 0$  und wir beobachten  $n_{\text{obs}} = 0$ . Für  $1-\gamma = 0.95$ ,

$$\gamma = P(n = 0; s, b = 0) = e^{-s} \rightarrow s_{\text{up}} = -\ln \gamma \approx 3.00$$

# Berechnung von Poisson-Limits

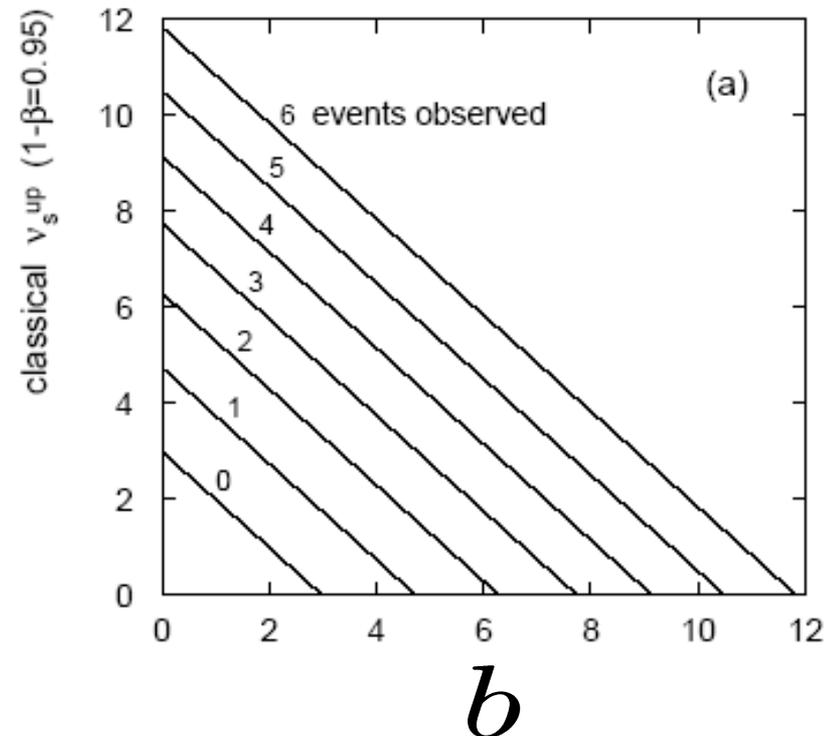
Um  $s_{lo}$ ,  $s_{up}$ , zu bestimmen, können wir die Beziehung zur  $\chi^2$ -WDF nutzen:

$$s_{lo} = \frac{1}{2} F_{\chi^2}^{-1}(\alpha; 2n) - b$$

Quantile der  $\chi^2$ -WDF

$$s_{up} = \frac{1}{2} F_{\chi^2}^{-1}(1 - \beta; 2(n + 1)) - b$$

Für statistische Fluktuation der Anzahl beobachteter Ereignisse nach unten im Vergleich zur Untergrunderwartung kann dies ein negatives Ergebnis für  $s_{up}$  liefern; i.e. das Konfidenzintervall im phys. sinnvollen Bereich ist leer



# Limits nahe der Physikalischen Grenze

Annahme z.B.  $b = 2.5$  und wir beobachten  $n = 0$ .

Wenn wir das  $CL = 0.9$  wählen, find wir aus der Formel für  $s_{\text{up}}$

$$s_{\text{up}} = -0.197 \quad (CL = 0.90)$$

Physiker (eventuell):

Wir wissen bereits dass  $s \geq 0$  gilt bevor wir gemessen haben; wir können kein negatives oberes Limit angeben um das Ergebnis unseres teuren Experimentes zu präsentieren!

Statistiker (Frequentist):

Das Intervall ist so konstruiert, dass es den wahren Wert nur in 90% der Fälle abdeckt — dies war eben einer von den 10% anderen Fällen.

Keine außergewöhnliche Situation, wenn Grenzwert/Ausschlussgrenze für Parameter nahe am physikalisch sinnvollen Bereich liegt

# Erwartetes Limit für $s=0$

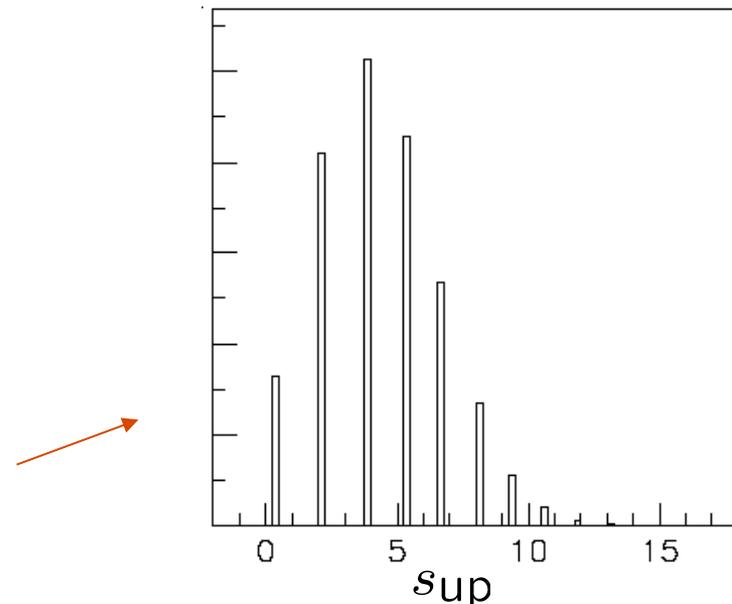
Physiker: Ich hätte  $CL = 0.95$  verwenden sollen, dann  $s_{up} = 0.496$

Oder besser noch: für  $CL = 0.917923$  erhält man  $s_{up} = 10^{-4}$ !

einfacher Vernünftigkeitstest: für  $b = 2.5$ , typische Poisson-Fluktuation in  $n$  ist mindestens  $\sqrt{2.5} = 1.6$ . Wie kann das Limit so klein sein?

Betrachte den Modalwert der Limits für die Hypothese “Nur-Untergrund” ( $s = 0$ )  $\rightarrow$  erwartete Sensitivität

Verteilung der 95% CL-Limits für  $b = 2.5$ ,  $s = 0$ .  
Mittelwert = 4.44



# Bayesianische Glaubwürdigkeitsintervalle

In der Bayesianischen Statistik benötigen wir eine A-Priori WDF  $\pi(\theta)$ . Diese reflektiert unseren Grad an Glauben über den Parameter  $\theta$  bevor wir das Experiment durchführen.

Bayes Theorem sagt uns wie unserem “Glauben” im Licht der Daten/ Messung abändern sollen:

$$p(\theta|x) = \frac{L(x|\theta)\pi(\theta)}{\int L(x|\theta')\pi(\theta') d\theta'} \propto L(x|\theta)\pi(\theta)$$

Integriere Posterior-WDF  $p(\theta | x)$  um das Glaubwürdigkeitsintervall mit beliebigem Wahrscheinlichkeitsinhalt zu erhalten.

Z.B. erhalten wir für einen Poisson-Parameter  $s$  die obere Grenze zu 95% CL aus der Gleichung:

$$0.95 = \int_{-\infty}^{\text{sup}} p(s|n) ds$$

Für Bayesianer steckt alle Information in der Likelihoodfunktion aus der aktuellen Messung. Die Wiederholbarkeit ist für ihn irrelevant.

# A-Priori-WDF für Poisson-Parameter

Berücksichtige unsere Kenntnis, dass  $s \geq 0$  sein muss, durch A-Priori-WDF  $\pi(s) = 0$  for  $s < 0$ . Oft wird versucht dadurch die A-Priori-Unkenntnis/ Ignoranz zu beschreiben, z.B

$$\pi(s) = \begin{cases} 1 & s \geq 0 \\ 0 & \text{otherwise} \end{cases}$$

Diese WDF ist nicht korrekt normiert, aber das ist in Ordnung so lange die Likelihoodfunktion für große Werte von  $s$  verschwindet.

Die WDF ist nicht invariant unter Transformation des Parameters. Wenn wir eine gleichförmige Verteilung in der Masse eines neuen Teilchens verwendet hätten, dessen Produktionsrate mit der Masse abfällt, dann wäre das eine nicht gleichförmige WDF in der Anzahl der erwarteten Signalereignisse.

Der obige Ansatz reflektiert nicht wirklich einen vernünftigen Ansatz für unseren Grad an Glauben, aber wird dennoch oft als Referenzpunkt verwendet. Oder wird als Rezept für Bildung von KI angesehen, dessen frequentistische Eigenschaften untersucht werden können (Abdeckungswahrscheinlichkeit abhängig von wahren  $s$ )

# Bay. KI mit gleichförmiger A-Priori-WDF für s

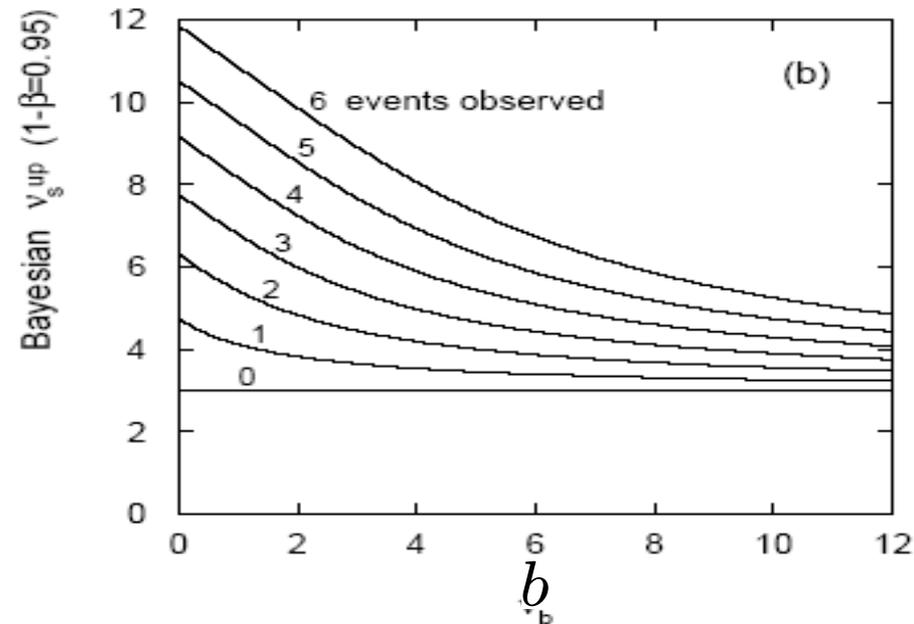
Löse das Problem numerisch um die obere Grenze  $s_{\text{up}}$  zu bestimmen.

$$L(n|s) = \frac{(s+b)^n}{n!} e^{-(s+b)} \quad 1 - \alpha = \int_{-\infty}^{s_{\text{up}}} p(s|n) ds = \frac{\int_{-\infty}^{s_{\text{up}}} L(n|s) \pi(s) ds}{\int_{-\infty}^{\infty} L(n|s) \pi(s) ds} \quad \pi(s) = \begin{cases} 0 & s < 0 \\ 1 & s \geq 0 \end{cases} \quad \alpha = e^{-s_{\text{up}}} \frac{\sum_{m=0}^n (s_{\text{up}} + b)^m / m!}{\sum_{m=0}^n b^m / m!}$$

Für den Spezialfall  $b = 0$ , ist die Bayesianische obere Ausschlussgrenze mit flacher A-Priori-WDF in  $s$  numerisch äquivalent zu klassischen frequentistischen Limits.

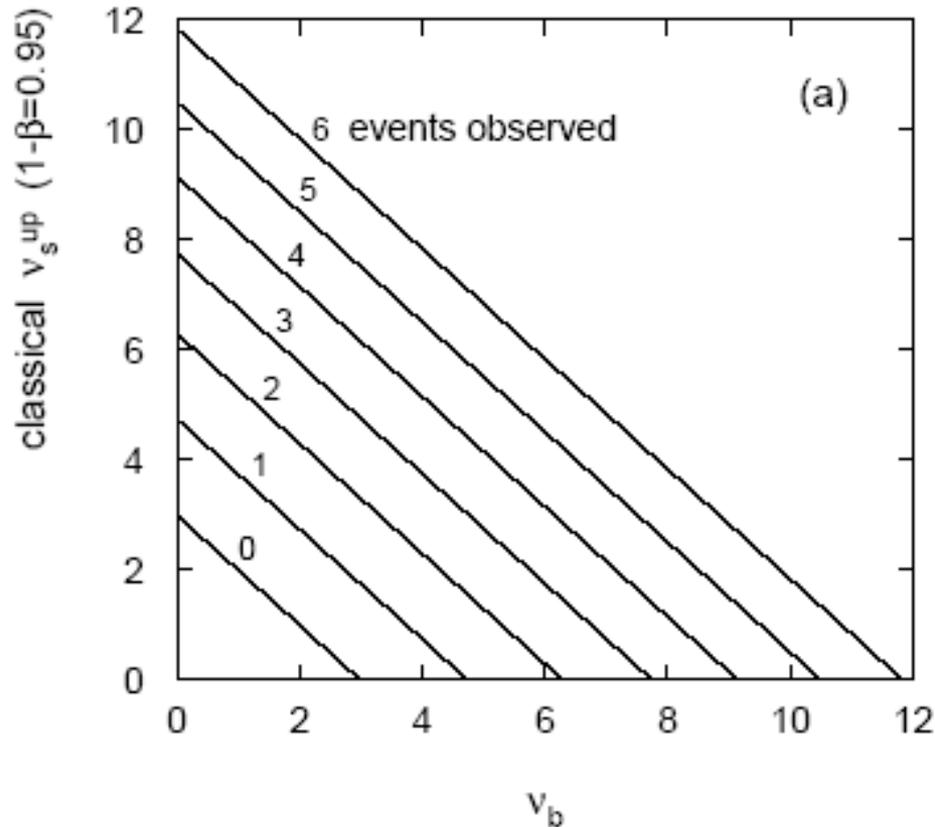
Für andere  $b$ -Werte ist das Bayesianische Limit immer größer als das klassische ("konservativ"). D.h. Abdeckungswkt. größer als angegebenes CL

Niemals negative obere Grenzen bzw. leere Konfidenzintervalle  
Unabhängig von  $b$  wenn  $n=0$ .



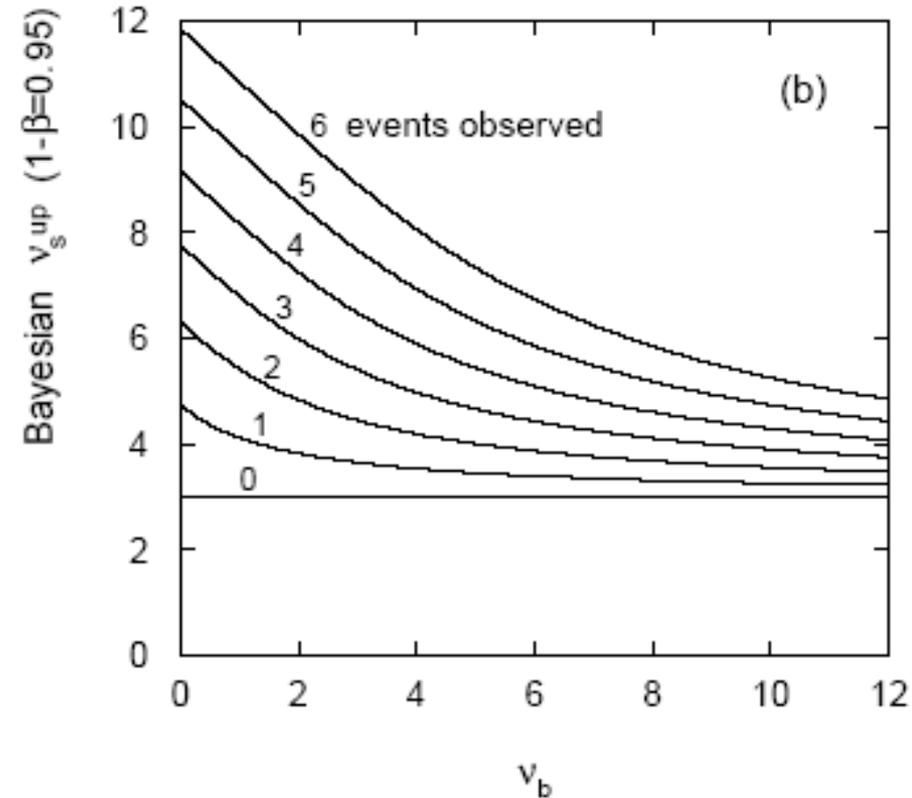
# Vgl. Klassische und Bayesianische Grenze (CL=95%)

## Klassisches Limit



Obere Grenze kann „0“ sein  
z.b: 0 beobachtet für  $b=3$

## Bayesianisches Limit

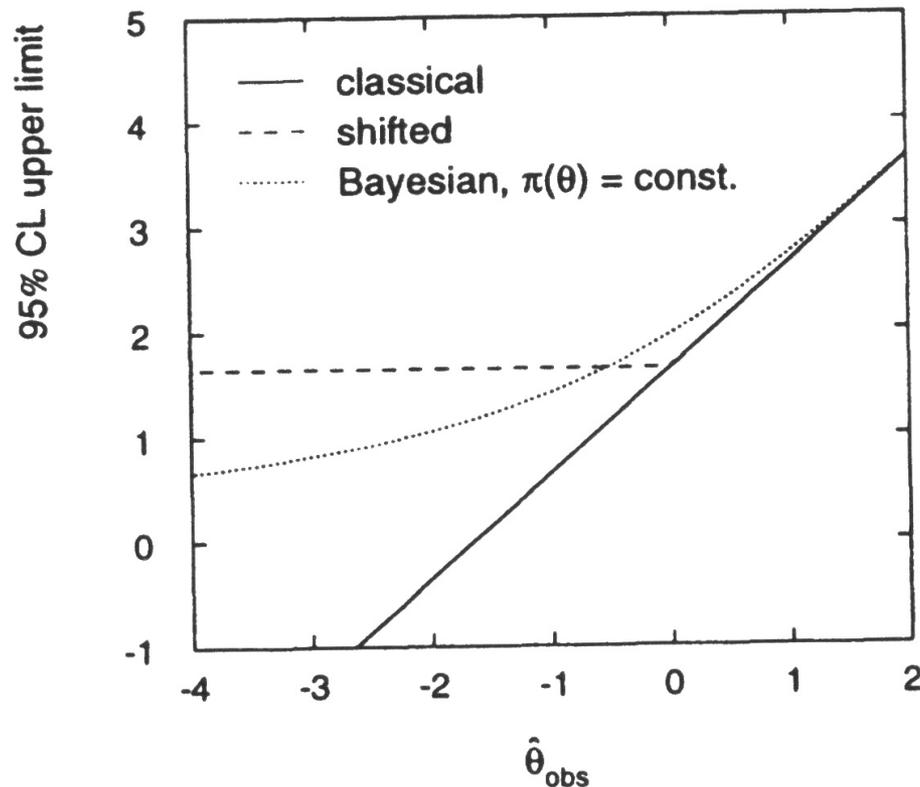


Obere Grenze immer  
größer oder gleich „3“

# Vergleich verschiedener Grenzen: Gauss-WDF

Annahme: WDF für Schätzer ist Standardgauss. i.e. mit Varianz=1  
physikalischer Bereich für Schätzer  $\theta \geq 0$ .

Weitere Möglichkeit für  $\theta_{\text{obs}} < 0$ : setze Schätzer auf  $\theta=0$



**Fig. 9.8** Upper limits at 95% confidence level for the example of Section 9.8 using the classical, shifted and Bayesian techniques. The shifted and classical techniques are equal for  $\hat{\theta}_{\text{obs}} \geq 0$ .

# Das Flip-Flop-Problem für frequentistische KI

Im Prinzip vor Messung festlegen, ob man einseitige oder zweiseitige Konfidenzintervalle bestimmen wird.

Annahme: Parameter = 0 ohne neue Physik

In der Praxis oft: wenn zweiseitiges 95%-KI die Null nicht enthält, dann wird zweiseitiges KI zu 68% CL angegeben, sonst wird obere Grenze zu 95% CL angegeben.

Problem: Abdeckungswahrscheinlichkeit wird für bestimmte Werte des wahren Parameters kleiner als das angegebene Konfidenzniveau

Lösung: sogenannter “Unified Approach”/Cousins-Feldman-Intervalle

Neuentdeckt für HEP in 1998. Basiert auf Likelihoodquotienten.

Details im nächsten Semester

# Zusammenfassung Konfidenzintervalle

Im Grenzfall großer Stichproben und nicht in der Nähe von physikalischer Grenze ist “ $\pm 1$  Standardabweichung” alles was braucht für 1-dim. Konfidenzintervalle zu 68% CL.

Frequentistische Konfidenzintervalle:

Komplizierte Konstruktion! Intervall ist Zufallsvariable die den wahren Wert mit festgesetzter Wkt bei Wiederholung der Messung beinhaltet.

Kann aus Invertierung eines Hypothesentests gewonnen werden; übriggebliebene Freiheit ist, welchen Test man wählt.

Log-Likelihood-Funtion kann verwendet werden um approximative Konfidenzintervalle zu bestimmen (in 1- oder mehr Dimensionen)

Flip-Flop-Problem und leere Intervalle: durch Likelihoodquotienten lösbar

Bayesianisches Konfidenzintervalle:

Konzeptuell einfach: einfach nur Posterior-WDF integrieren.

Verlangt aber die Wahl einer A-Priori-WDF. Keine leeren Intervalle

Abdeckung nicht festgelegt durch Konstruktion, aber nicht Ziel der Methode.