

Moderne Methoden der Datenanalyse

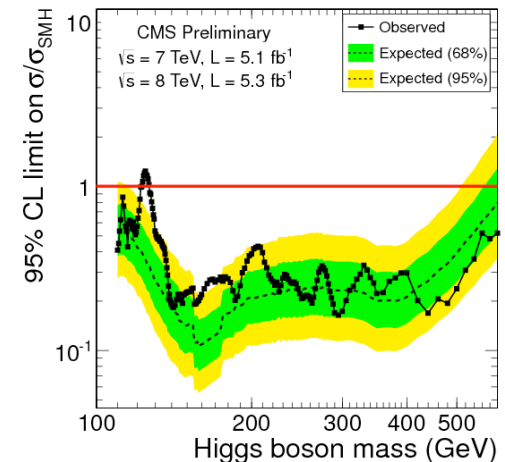
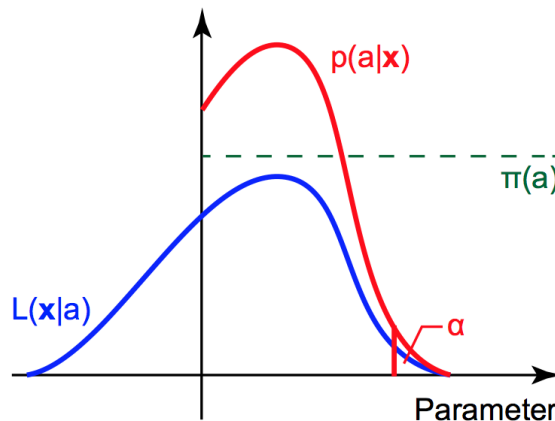
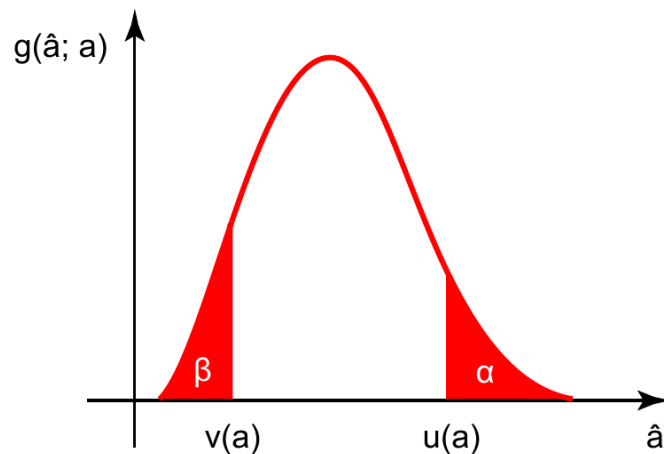
Zusammenfassung Konfidenz-Intervalle (1)

Andreas B. Meyer



13. Juni 2017

Fakultät für Physik
Institut für Experimentelle Kernphysik - IEKP



■ Frequentistische Sichtweise: Es existiert ein fester wahrer Wert a

- Wahre Werte sind wahr, haben also keine Unsicherheit
- Wahrscheinlichkeitsbegriff bezogen auf prinzipiell messbare Häufigkeit (z.B. QM)

Das Konfidenz-Intervall zu einem Konfidenz-Niveau $CL = p\%$ deckt den wahren Wert a in $p\%$ aller Fälle ab.

- Wahrscheinlichkeitsaussagen nur über das Intervall, nicht den wahren Wert
- “Inversion” durch Neyman-Konstruktion, Abdeckung “by design”

■ Bayes'sche Sichtweise

- Der wahre Wert a hängt von Voraussetzungen ab.
- “Prior” beschreibt degree of belief, dass a bestimmte Werte annehmen kann.
- Wahrer Wert besitzt eine Unsicherheit, die von der Messung abhängig ist

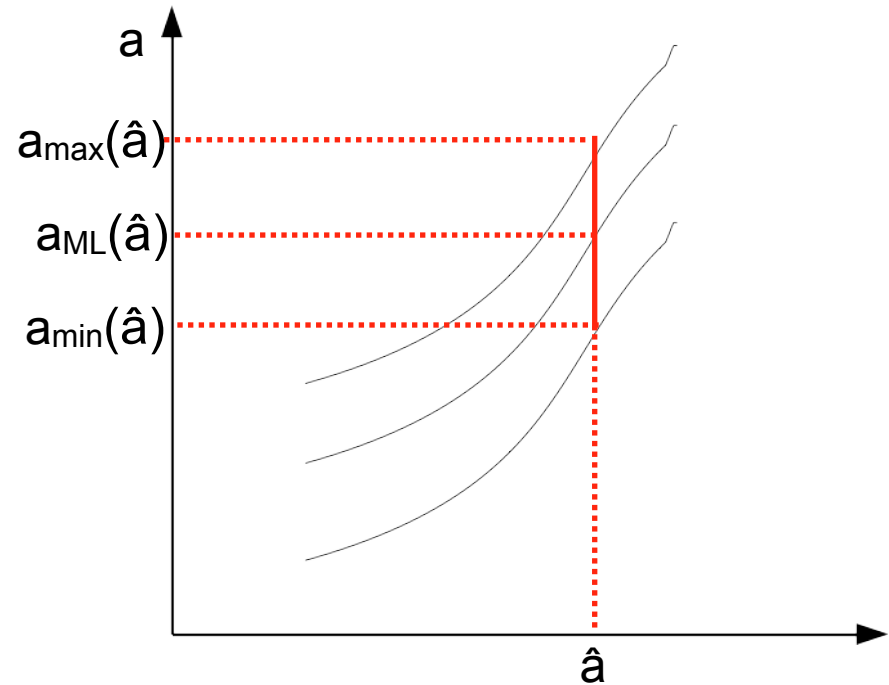
Konstruiere Wahrscheinlichkeitsdichte für unbekannten Wert a

$$f(a|\hat{a}) \propto \mathcal{L}(\hat{a}|a) \cdot \pi(a)$$

- Posterior $f(a|\hat{a})$ benötigt Likelihood-Funktion $\mathcal{L}(\hat{a}|a)$ und Prior $\pi(a)$
- Abdeckung muss explizit überprüft werden (z.B. mit toy-MC)

Neyman-Konstruktion

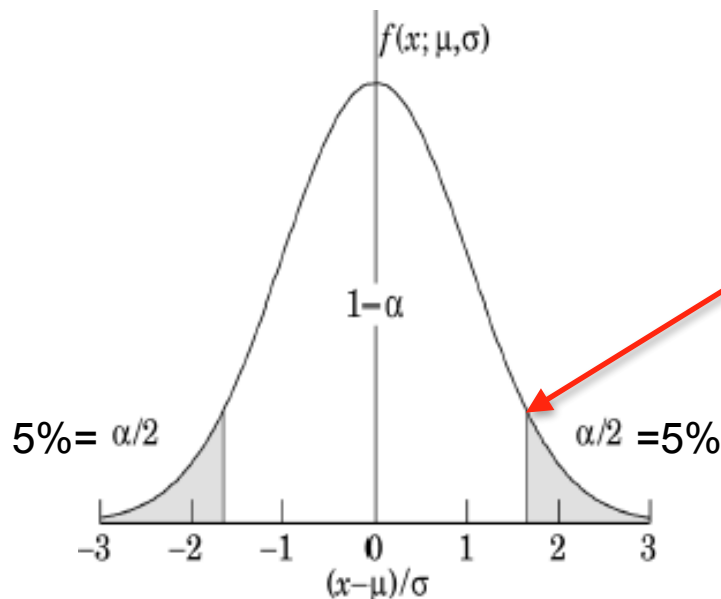
- Die Funktionen $u(a)$ und $v(a)$ beschreiben die Intervall-Grenzen für ein Konfidenz-Niveau als Funktion eines hypothetischen wahren Werts
- Umgekehrt liefern die Umkehrfunktionen von $u(a)$ und $v(a)$ die Intervall-Grenzen $a_{\max}(\hat{a}) = u^{-1}(\hat{a})$ und $a_{\min}(\hat{a}) = v^{-1}(\hat{a})$ ausgehend vom Messwert $\hat{a}$
- Bedingung: $u(a)$ und $v(a)$ monoton und Umkehrfunktionen existieren
- Hier ist jetzt auch der wahrscheinlichste Wert für a eingezeichnet. Streng frequentistisch hat a aber keine Unsicherheit



- Systematische Vorschrift aus frequentistischer Sicht zur
“Inversion der Wahrscheinlichkeitsaussagen”:
 - Gegeben ein wahrer Wert a , die Wahrscheinlichkeit $\hat{a}$ zu messen ist ...
 - Gegeben eine Messung $\hat{a}$, die Wahrscheinlichkeit, dass wahrer Wert a im Intervall liegt, ist ...
- Aussage über Abdeckung des wahren Werts durch das Intervall, nicht über Wahrscheinlichkeit des wahren Werts selbst
- Neyman-Konstruktion garantiert exakte Abdeckung (gemäß Vorgabe) und Überabdeckung für diskrete Verteilungen
 - Zentrale Konfidenz-Intervalle → Messungen
 - Einseitige Konfidenz-Intervalle → obere oder untere Schranken
 - Unified Frequentist Approach: verwendet Likelihood-Ratio als Ordnungsprinzip (Vermeidung von “Flip-flopping” und “leeren Intervallen”)
- Bemerkung: Betrachte Vorschriften immer im Verhältnis zur relativen Größe der statistischen Effekte und zum Aufwand. Andere Fehler als statistisch-methodische, z.B. systematische Unsicherheiten, spielen bei Messungen meist eine große Rolle

Intervall-Schätzung

- Üblich sind CL=68.3%, 90% oder 95%, oder entsprechende Quantile der Gauß-Verteilung
- Zweiseitige Intervalle: Messungen
- Einseitige Intervalle: Obere oder untere Schranken, z.B. bei Suchen



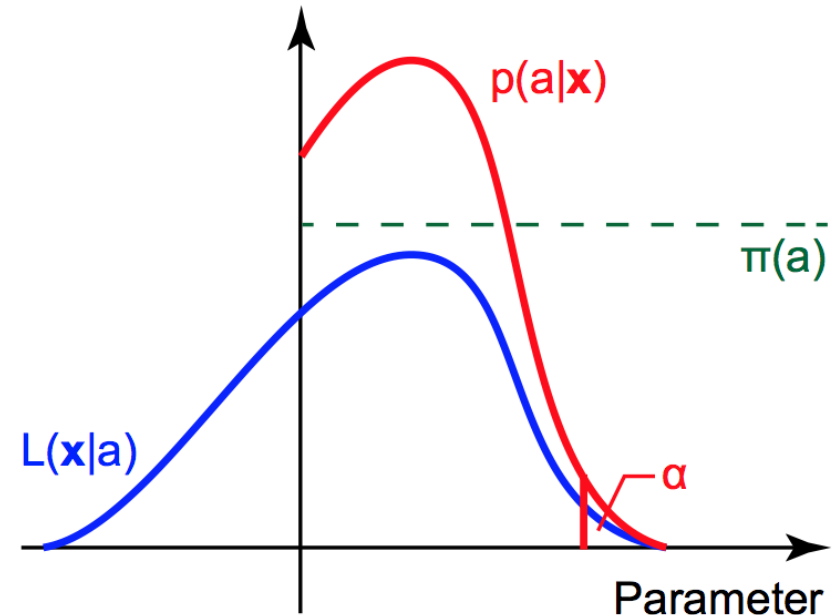
Root: TMath::ChisquareQuantile

$\alpha_{\text{zweiseitig}}(\delta)$

$1 - \alpha$	α	δ
0.683	0.317	1σ
0.90	0.1	1.65σ
0.95	0.05	1.96σ
0.9545	0.0455	2σ
0.9973	0.0027	3σ
	5.7×10^{-7}	5σ

$$\alpha_{\text{einseitig}}(\delta) = \alpha_{\text{zweiseitig}}(\delta)/2$$

- Konfidenzintervall aus Posterior-Verteilung $p(a|x)$
- Wahl der Prior-Verteilung $\pi(a)$ beinhaltet physikalisches Urteil
- Physikalische Grenzen sind in Prior-Verteilung berücksichtigt z.B. falls a immer positiv.
- Frequentistische Abdeckung nicht mehr automatisch garantiert



1. Umnormierung der Wahrscheinlichkeit im physikalischen Bereich
2. Konfidenzintervall durch Integration

Poisson-Signal und Untergrund: mit Bayes' Prior

- Betrachte typische Suche, d.h. Anzahl der Signalereignisse ν_s klein
- Signal + Untergrund ist Poisson-verteilt: $p(n | \nu_s, \nu_b) = p(n | \nu = \nu_s + \nu_b)$
- Bestimmung ν und Subtraktion von ν_b liefert Schätzwert $\hat{\nu}_s = n - \nu_b$
- Oberes Limit (95% CL) für ν_s als Funktion des erwarteten U.-grunds ν_b , für verschiedene n_{obs} :

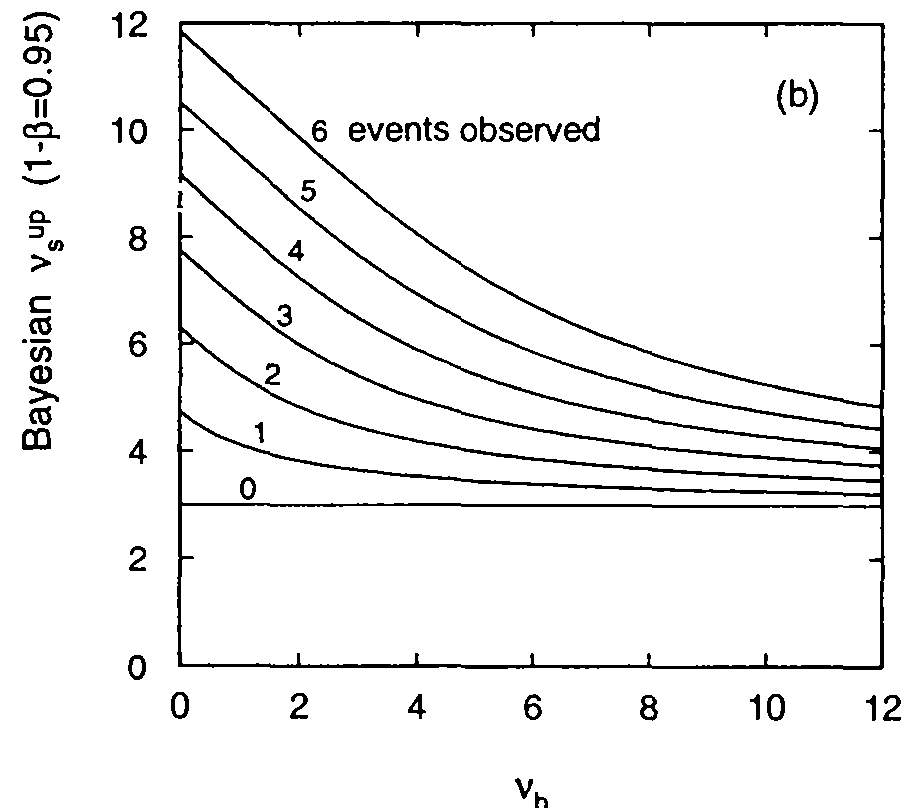
■ Bayes' Prior:

$$\pi(\nu_s < 0) = 0 \text{ and } \pi(\nu_s \geq 0) = \text{const}$$

liefert gute Eigenschaften:

- gleiches Ergebnis für $\nu_b = 0$
- höhere Grenze für $\nu_b > 0$

d.h. Abdeckung ist hinreichend



Modifizierter (Frequentistischer) Ansatz: CL_s

- Gesucht: Robuste obere Schranke (oder Messung) eines hypothetischen Signals (gemessen mit Observabler d), aber ohne Bayes' Prior
- Bestimme obere Grenzen für zwei Hypothesen:
 - H_1 : Gemessene Ereignisse bestehen aus Untergrund und Signal:
 $d = s + b$ → liefert p-Wert = " CL_{b+s} "
 - H_0 : Gemessene Ereignisse bestehen nur aus Untergrund:
 $d = b$, d.h. $s=0$ → liefert p-Wert = " CL_b "

$$CL_s = \frac{CL_{s+b}}{CL_b} = \frac{\sum_{k=0}^d P(k; s + b)}{\sum_{k=0}^d P(k; b)} = \frac{p_\mu}{1 - p_b}$$

- CL_s renormiert gemessene Grenze auf die Abschätzung des Untergrunds:
 - Quantitativ ähnlicher Effekt wie mit konstantem Bayes' Prior
 - Vergleichbarkeit von Experimenten mit unterschiedlichem Untergrund
 - CL_s immer größer als p_μ → Abdeckung ok

G.Cowan, PDG,
Abschnitt 39.4.2.4

Zusammenfassung Konfidenz-Intervalle (1)

- Bestimme Intervall, in dem wahrer Wert mit vorgegebener Wahrscheinlichkeit liegt
- Relevant für nicht-Gauß-verteilte PDF

- Frequentistischer Ansatz:
 - Neyman-Konstruktion - Inversion der Wahrscheinlichkeitsverteilungen
 - Korrekte Abdeckung per Definition
 - Unified Frequentist Approach: verwendet Likelihood-Ratio als Ordnungsprinzip (Vermeidung von Flip-flopping und leeren Intervallen)

- Bayes'sche Priors:
 - Zur Vermeidung unphysikalischer Ergebnisse
 - Abdeckung muss separat getestet werden

- Modifizierter frequentistischer Ansatz CL_s :
 - Etablierte, robuste Methode, um statistische Fluktuationen des Untergrunds zu berücksichtigen
 - Überabdeckung per Definition