

Einführung in die Statistik und Datenanalyse (5)

Martin Völkl
martin.voelkl@uni-tuebingen.de

Universität Tübingen
2017-03-20

- 1 Einführung und systematische Unsicherheiten
- 2 Werkzeugkiste
- 3 Wahrscheinlichkeit und Unsicherheit
- 4 Frequentistische Methoden
- 5 **Bayesianische Methoden**
 - Bayes' Theorem
 - Störparameter
 - Einfache Monte Carlo–Methoden
 - Markov–Chain Monte Carlo

$$p(A|B) = \frac{p(B|A) \cdot p(A)}{\sum_A p(B|A_i) \cdot p(A_i)}$$

- Für Parameter wie in der vorherigen Vorlesung:

$$p(\mu|\vec{d}) = \frac{L(\mu|\vec{d}) \cdot p(\mu)}{\int_{\mu} L(\mu|\vec{d}) \cdot p(\mu) d\mu}$$

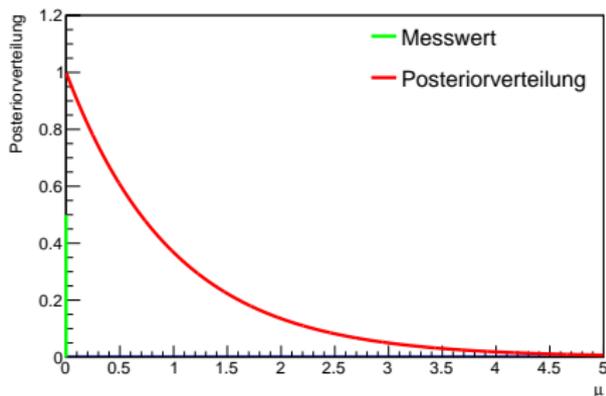
- Für Hypothesen H_0 , H_1

$$p(H_1|\vec{d}) = \frac{p(\vec{d}|H_1) \cdot p(H_1)}{p(\vec{d}|H_1) \cdot p(H_1) + p(\vec{d}|H_0) \cdot p(H_0)}$$

- Hypothesen und Parameter haben Wahrscheinlichkeiten!
- Nenner ist "nur" Normierung

Beispiel: Poisson-Messung

- Für Messung von Zahlen aus Poisson-Prozess $n \pm \sqrt{n}$ nur gültig für große n
- Frequentistisch: Konfidenzintervalle \rightarrow schwierig zu interpretieren



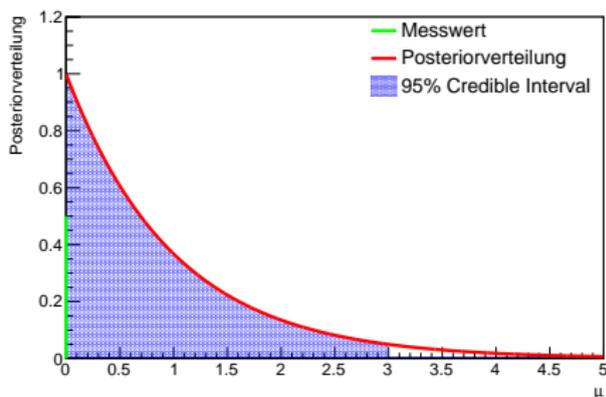
- Bayesianisch: Posteriorverteilung

$$p(\mu|\vec{d}) = \frac{L(\mu|\vec{d}) \cdot p(\mu)}{\int_{\mu} L(\mu|\vec{d}) \cdot p(\mu) d\mu} = \exp(-x)$$

- Annahme flacher Prior $p(\mu) = 1$

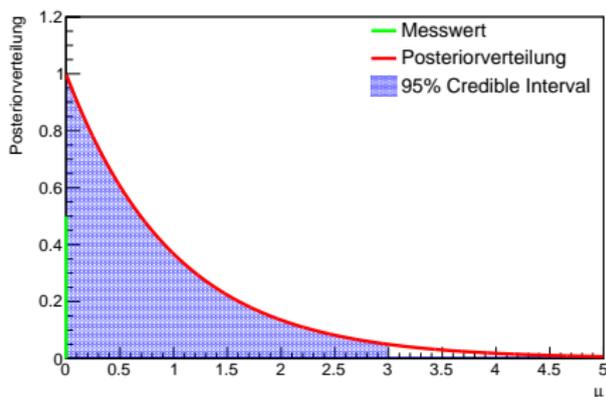
Beispiel: Poisson-Messung (2)

- Fläche unter der Kurve ist Wahrscheinlichkeitsverteilung für Parameter
- Einfache Interpretation
- Statt Konfidenzintervall: Credible interval
- 95% credible interval: Der Parameter ist mit einer Wahrscheinlichkeit von 95% im Intervall
- Keine Schätzfunktionen, p-Werte oder sonstiges
- Man muss "nur" Integrieren können



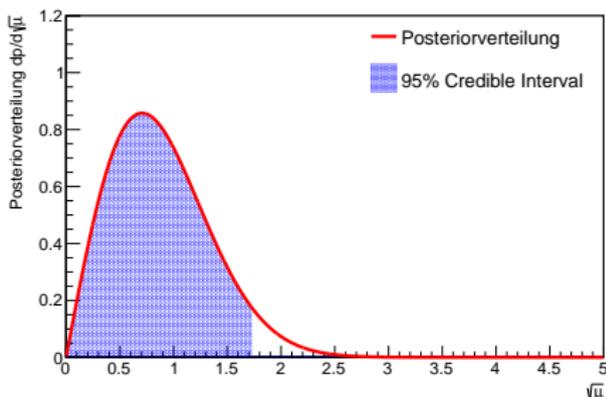
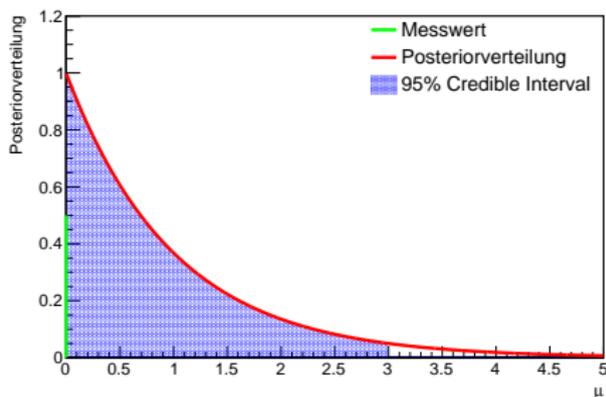
Beispiel: Poisson-Messung (2)

- Fläche unter der Kurve ist Wahrscheinlichkeitsverteilung für Parameter
- Einfache Interpretation
- Statt Konfidenzintervall: Credible interval
- 95% credible interval: Der Parameter ist mit einer Wahrscheinlichkeit von 95% im Intervall
- Keine Schätzfunktionen, p-Werte oder sonstiges
- Man muss "nur" Integrieren können

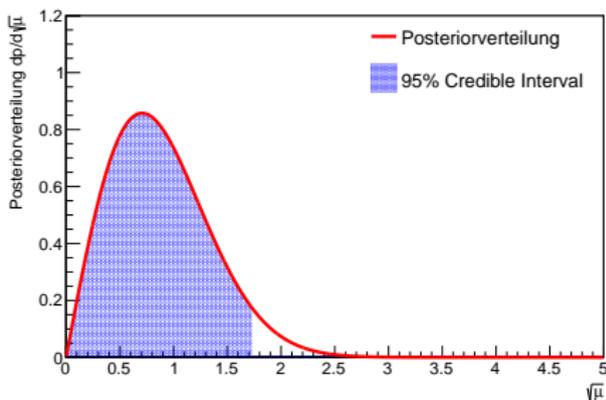
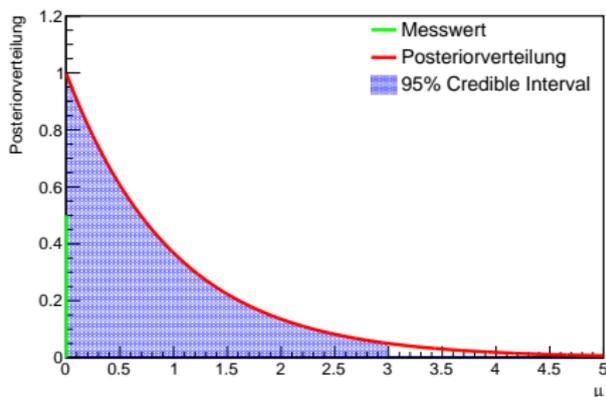


Beispiel: Fehlerfortpflanzung

- Kennen Posterior für Variable μ
- Was ist Posterior für $f(\mu)$?
- Extrem simpel bei Bayes':
Propagiere
Wahrscheinlichkeitsverteilung
- Kein Bias o. Ä.
- keine Näherungen notwendig

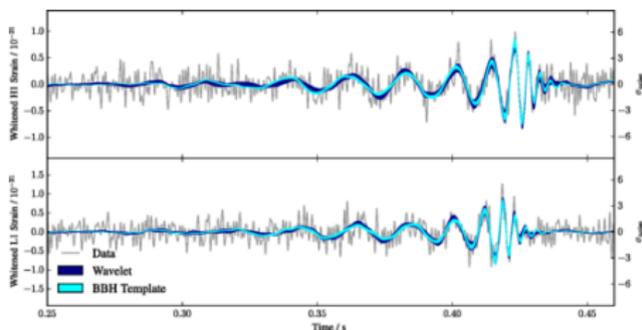


- Priors geben Wissen über Parameter vor der Messung an
- Sollten alles Wissen enthalten
- Für kontinuierliche Parameter oft nicht einfach, e.g. konstanter Prior von μ ist nicht konstant für μ^2
- Häufig verwendete Priors:
 - Flacher Prior $p(\mu) = \text{const.}$ für Ortsparameter
 - $p(\mu) = 1/\mu$ für Skalenparameter ($p(\log \mu) = \text{const.}$)
- Wenn Daten viele Informationen enthalten hat der Prior wenig Auswirkungen auf das Resultat



Gravitationswellenmessung

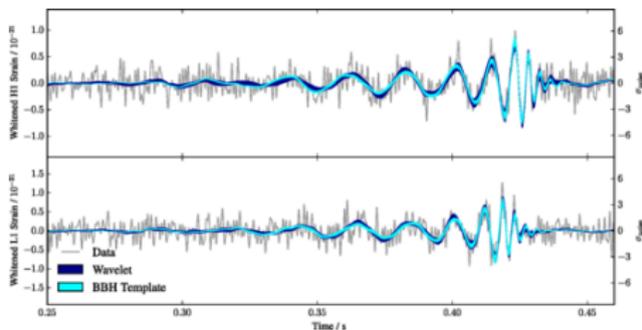
- Unbekannte Parameter:
Massen der schwarzen Löcher,
Drehimpulse, Richtung des
Systems, Entfernung,
Ausrichtung zur Erde,
Exzentrizität, ...
- Insgesamt 17 unbekannte
Parameter
- Für jedes Parameterset ist
Likelihood bekannt (im Prinzip
Normalverteilung um das
erwartete Signal)
- Posterior $p(\mu_1, \mu_2, \dots | \vec{d})$ hat
17 Parameter, wie bekommt
man z.B. $p(\mu_3 | \vec{d})$?



	EOBNR	IMRPhenom	Overall
Detector-frame total mass M/M_{\odot}	$70.3^{+2.3}_{-4.8}$	$70.9^{+4.0}_{-3.0}$	$70.6^{+0.0+0.5}_{-2.1+1.3}$
Detector-frame chirp mass M_c/M_{\odot}	$30.2^{+2.5}_{-3.0}$	$30.6^{+1.8}_{-1.8}$	$30.4^{+2.1+0.2}_{-1.0+0.2}$
Detector-frame primary mass m_1/M_{\odot}	$39.4^{+5.5}_{-4.0}$	$38.5^{+5.6}_{-3.6}$	$38.9^{+5.6+0.6}_{-4.5+0.4}$
Detector-frame secondary mass m_2/M_{\odot}	$30.9^{+4.8}_{-4.4}$	$32.2^{+3.9}_{-3.6}$	$31.6^{+2.0+1.1}_{-2.1+0.9}$
Detector-frame final mass M_f/M_{\odot}	$67.1^{+4.0}_{-4.4}$	$67.6^{+3.9}_{-3.5}$	$67.4^{+4.1+0.4}_{-4.0+0.2}$
Source-frame total mass M^{source}/M_{\odot}	$65.0^{+5.0}_{-4.4}$	$65.0^{+4.0}_{-3.6}$	$65.0^{+3.4+0.8}_{-2.0+0.7}$
Source-frame chirp mass M_c^{source}/M_{\odot}	$27.9^{+2.3}_{-2.8}$	$28.1^{+1.7}_{-1.6}$	$28.0^{+2.0+0.3}_{-1.0+0.1}$
Source-frame primary mass m_1^{source}/M_{\odot}	$36.3^{+5.3}_{-4.5}$	$35.3^{+5.2}_{-3.4}$	$35.8^{+5.3+0.9}_{-4.0+0.1}$
Source-frame secondary mass m_2^{source}/M_{\odot}	$28.6^{+4.4}_{-4.0}$	$29.6^{+3.5}_{-3.1}$	$29.1^{+2.8+0.1}_{-2.4+0.1}$
Source-frame final mass M_f^{source}/M_{\odot}	$62.0^{+4.4}_{-4.0}$	$62.0^{+3.7}_{-3.1}$	$62.0^{+4.1+0.7}_{-2.7+0.4}$
Mass ratio q	$0.79^{+0.18}_{-0.18}$	$0.84^{+0.14}_{-0.20}$	$0.82^{+0.17+0.05}_{-0.20+0.05}$
Effective inspiral spin parameter χ_{eff}	$-0.09^{+0.19}_{-0.17}$	$-0.05^{+0.13}_{-0.15}$	$-0.07^{+0.16+0.04}_{-0.17+0.04}$
Dimensionless primary spin magnitude a_1	$0.32^{+0.45}_{-0.28}$	$0.32^{+0.53}_{-0.29}$	$0.32^{+0.49+0.06}_{-0.29+0.01}$
Dimensionless secondary spin magnitude a_2	$0.57^{+0.40}_{-0.19}$	$0.34^{+0.54}_{-0.04}$	$0.44^{+0.50+0.08}_{-0.40+0.02}$
Final spin a_f	$0.67^{+0.09}_{-0.09}$	$0.66^{+0.04}_{-0.04}$	$0.67^{+0.05+0.02}_{-0.07+0.02}$
Luminosity distance D_L/Mpc	3001^{+170}_{-180}	4401^{+190}_{-180}	$4101^{+160+20}_{-180+20}$
Source redshift z	$0.083^{+0.035}_{-0.036}$	$0.093^{+0.029}_{-0.026}$	$0.088^{+0.032+0.005}_{-0.027+0.005}$
Upper bound on primary spin magnitude a_1	0.65	0.74	0.69 \pm 0.08
Upper bound on secondary spin magnitude a_2	0.93	0.78	0.89 \pm 0.13
Lower bound on mass ratio q	0.64	0.68	0.66 \pm 0.03
Log Bayes factor $\ln \mathcal{B}_{1/0}$	288.7 ± 0.2	290.3 ± 0.1	...

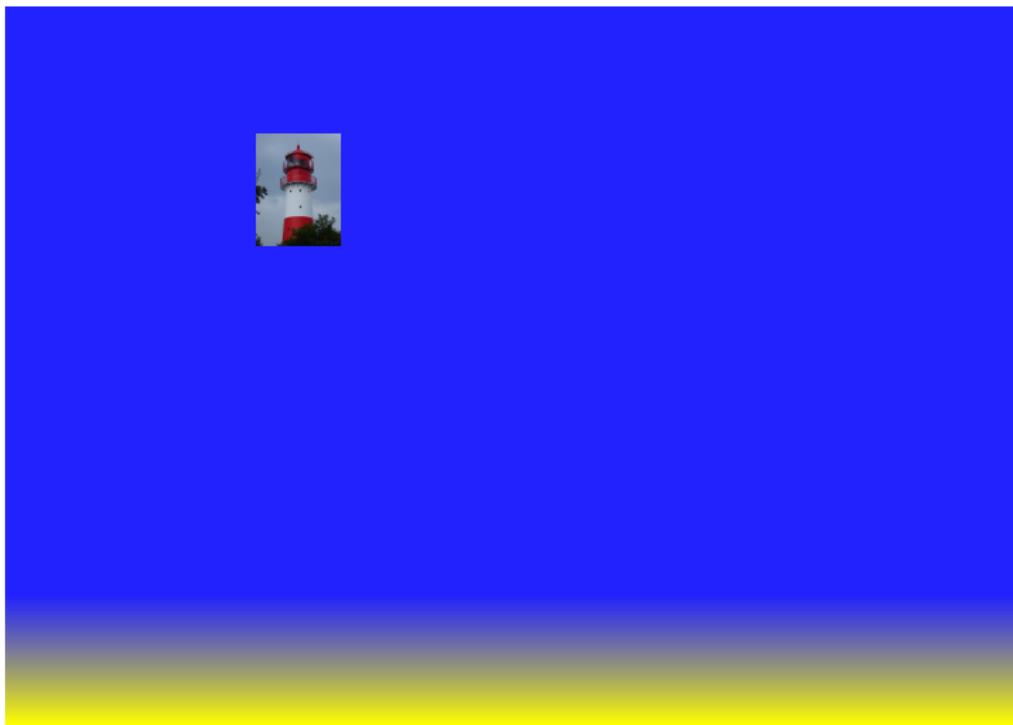
Gravitationswellenmessung (2)

- Posterior $p(\mu_1, \mu_2, \dots | \vec{d})$
- Wie bereits besprochen: Integrieren über alle unbenötigten Parameter
- Also: 16–dimensionales Integral lösen \rightarrow selbst numerisch problematisch
- *Bayesianische Statistik kann mit vielen Störparametern umgehen, braucht aber Methoden zur Integralberechnung*



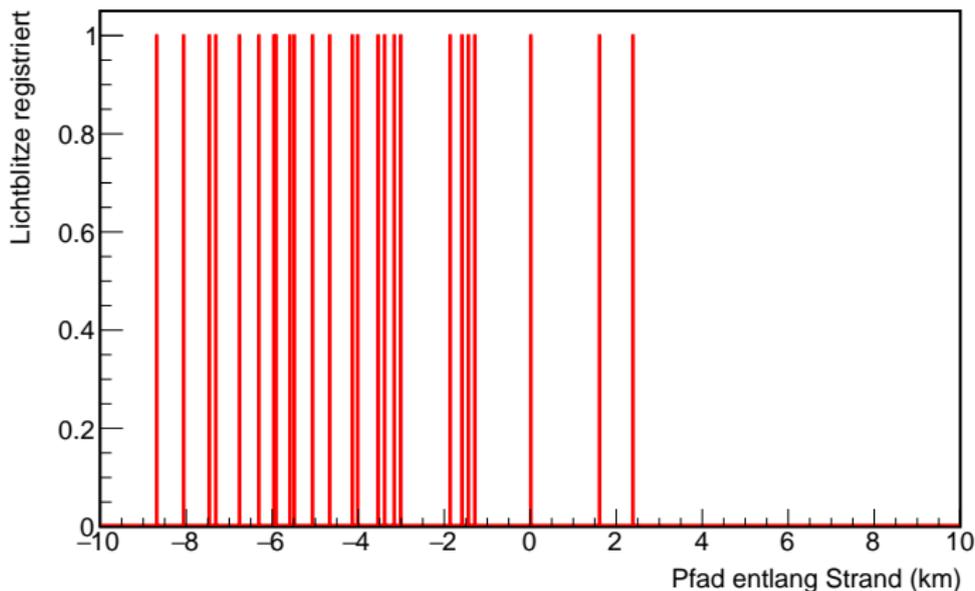
	EOBNR	IMRPhenom	Overall
Detector-frame total mass M/M_{\odot}	$70.3^{+5.7}_{-4.8}$	$70.9^{+4.0}_{-3.9}$	$70.6^{+4.8+0.8}_{-5.1+1.3}$
Detector-frame chirp mass \mathcal{M}/M_{\odot}	$30.2^{+2.5}_{-1.9}$	$30.6^{+1.8}_{-1.8}$	$30.4^{+1.8+0.2}_{-1.9+0.3}$
Detector-frame primary mass m_1/M_{\odot}	$39.4^{+5.5}_{-4.9}$	$38.5^{+5.6}_{-5.6}$	$38.9^{+5.6+0.6}_{-4.5+0.4}$
Detector-frame secondary mass m_2/M_{\odot}	$30.9^{+4.8}_{-4.4}$	$32.2^{+5.6}_{-4.8}$	$31.6^{+4.5+0.4}_{-4.7+0.9}$
Detector-frame final mass M_f/M_{\odot}	$67.1^{+4.4}_{-4.4}$	$67.6^{+5.5}_{-5.0}$	$67.4^{+4.4+0.4}_{-4.5+0.8}$
Source-frame total mass M^{source}/M_{\odot}	$65.0^{+4.4}_{-4.4}$	$65.0^{+5.6}_{-5.0}$	$65.0^{+4.0+0.3}_{-4.0+0.3}$
Source-frame chirp mass $\mathcal{M}^{source}/M_{\odot}$	$27.9^{+2.3}_{-1.8}$	$28.1^{+1.6}_{-1.6}$	$28.0^{+1.5+0.3}_{-1.5+0.3}$
Source-frame primary mass m_1^{source}/M_{\odot}	$36.3^{+5.5}_{-4.5}$	$35.3^{+5.2}_{-5.4}$	$35.8^{+5.1+0.9}_{-4.6+0.1}$
Source-frame secondary mass m_2^{source}/M_{\odot}	$28.6^{+4.4}_{-4.4}$	$29.6^{+4.5}_{-4.3}$	$29.1^{+4.4+0.1}_{-4.3+0.7}$
Source-frame final mass M_f^{source}/M_{\odot}	$62.0^{+4.0}_{-4.0}$	$62.0^{+4.4}_{-4.4}$	$62.0^{+3.7+0.4}_{-3.7+0.4}$
Mass ratio q	$0.79^{+0.18}_{-0.18}$	$0.84^{+0.14}_{-0.20}$	$0.82^{+0.12+0.01}_{-0.20+0.03}$
Effective inspiral spin parameter χ_{eff}	$-0.09^{+0.19}_{-0.05}$	$-0.05^{+0.15}_{-0.15}$	$-0.07^{+0.17+0.05}_{-0.17+0.05}$
Dimensionless primary spin magnitude a_1	$0.32^{+0.28}_{-0.28}$	$0.32^{+0.29}_{-0.29}$	$0.32^{+0.28+0.06}_{-0.28+0.01}$
Dimensionless secondary spin magnitude a_2	$0.57^{+0.40}_{-0.51}$	$0.34^{+0.31}_{-0.31}$	$0.44^{+0.40+0.08}_{-0.40+0.02}$
Final spin a_f	$0.67^{+0.08}_{-0.08}$	$0.66^{+0.08}_{-0.08}$	$0.67^{+0.07+0.02}_{-0.07+0.02}$
Luminosity distance D_L/Mpc	390^{+170}_{-180}	440^{+180}_{-190}	$410^{+160+40}_{-160+40}$
Source redshift z	$0.083^{+0.033}_{-0.036}$	$0.093^{+0.029}_{-0.036}$	$0.088^{+0.029+0.005}_{-0.029+0.008}$
Upper bound on primary spin magnitude a_1	0.65	0.74	0.69 ± 0.08
Upper bound on secondary spin magnitude a_2	0.93	0.78	0.89 ± 0.13
Lower bound on mass ratio q	0.64	0.68	0.66 ± 0.03
Log Bayes factor $\ln \mathcal{B}_{1,0}$	288.7 ± 0.2	290.3 ± 0.1	...

Das Leuchtturmproblem – Gemälde



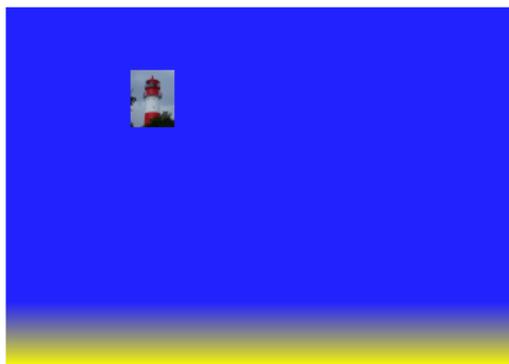
- Leuchtturm schießt Strahlen in zufällige Richtung (Photonen)
- Unbekannte Richtung, Position

Das Leuchtturmproblem – Messung



- Mit diesen 25 Lichtblitzen wollen wir Messen, wie weit der Leuchtturm entfernt ist

- Suchen Distanz von Strand R
- Position x_0 ist Störparameter
- Gemessen wird nur auf einer Länge von 20 km
- Prior $p(x_0, R) = \text{const.}$



- Ohne Herleitung, Wahrscheinlichkeit für einzelnen Lichtblitz i bei x_i :

$$p(x_i|x_0, R) = \frac{1}{R} \cdot \frac{1}{1 + \left(\frac{x_i - x_0}{R}\right)^2} \cdot \frac{1}{\arctan\left(\frac{x_{\max} - x_0}{R}\right) - \arctan\left(\frac{x_{\min} - x_0}{R}\right)}$$

- Mehrere Messungen: Produkt der einzelnen Wahrscheinlichkeiten
- 2-dimensional – hier "hochdimensional"
- Wie Posteriorverteilung berechnen/marginalisieren?

- Angenommen: Zufallszahlen (x, y) aus Verteilung $p(x, y)$ vorhanden
- Genügend Zahlen nähern Verteilung
- "Marginalisieren" → ignoriere andere Dimensionen
- Bei Bayes:

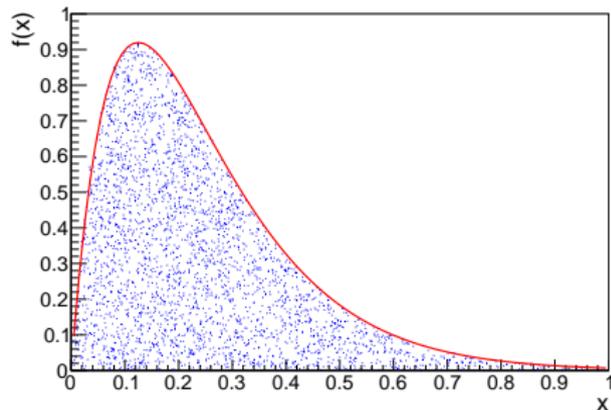
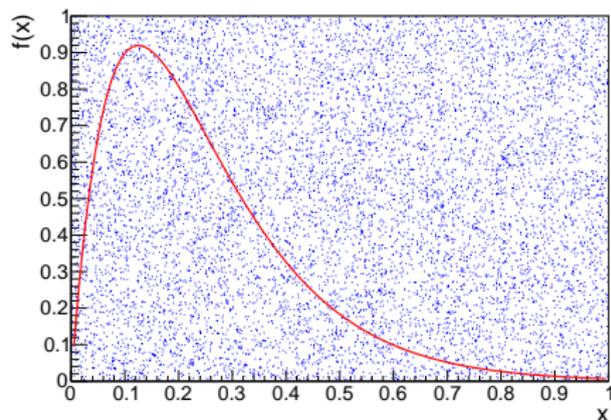
$$p(\vec{\mu}|\vec{d}) = \frac{L(\vec{\mu}|\vec{d}) \cdot p(\vec{\mu})}{\int L(\vec{\mu}|\vec{d}) \cdot p(\vec{\mu}) d\mu_1 \dots d\mu_n}$$

ist außerdem die Normierung störend

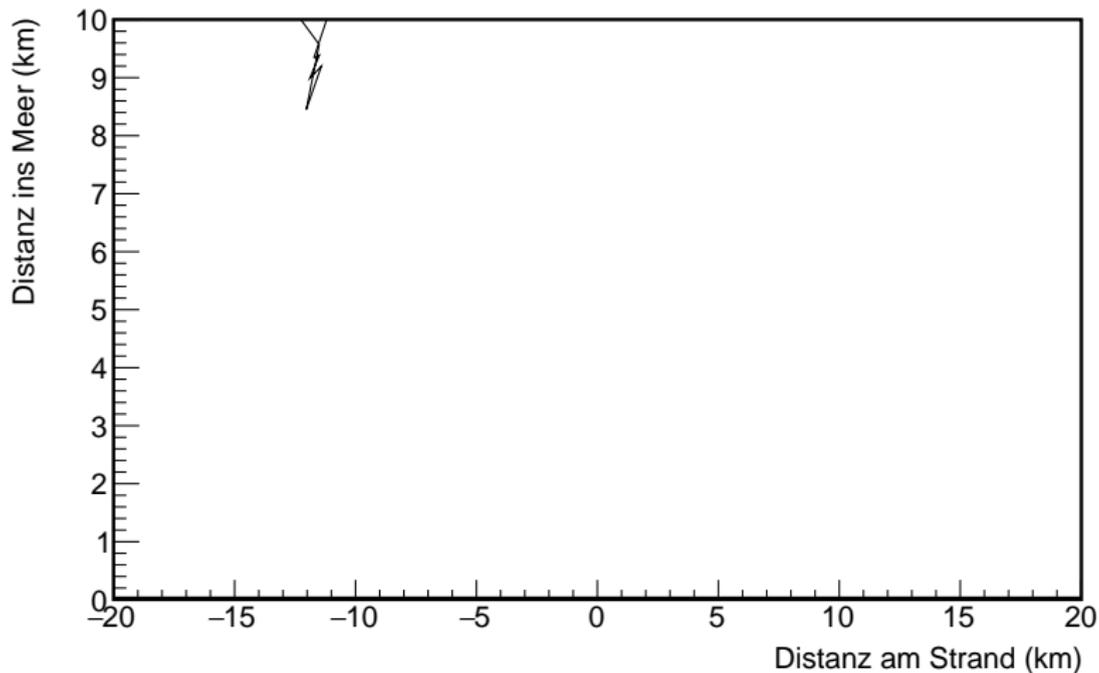
- Der Zähler ist für einen bestimmten Punkt leicht zu berechnen
- Wie kann man Zufallszahlen aus einer unbekanntem Verteilung ziehen; möglichst ohne die Normierung zu kennen?

Accept-Reject Monte Carlo

- Schnelle Zufallsgeneratoren bekannt
- Erzeuge zufällige Punkte in Rechteck, welches Funktion umschließt
- Behalte nur Punkte, die unterhalb der Funktion liegen
- Funktioniert unabhängig von der Normierung der Funktion
- Bei hoher Dimensionalität, viel

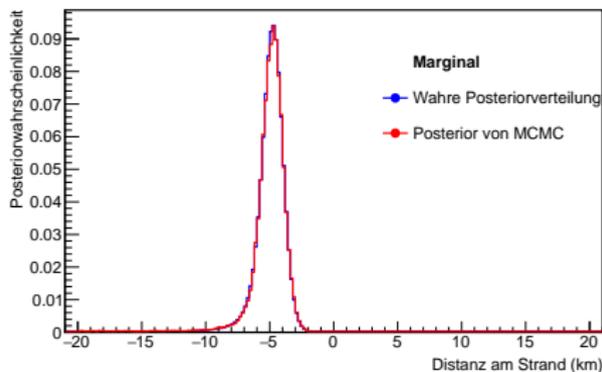


- Idee: Suche nächste Zufallszahl in der Nähe der letzten akzeptierten Zahl
- Beispiel Metropolis Algorithmus:
 - Von der momentanen Position \vec{x}_i , wähle einen zufälligen Schritt \vec{d} aus (führt nach $\vec{x}_i + \vec{d}$)
 - Vergleiche die Wahrscheinlichkeiten $p(\vec{x}_i)$ und $p(\vec{x}_i + \vec{d})$
 - Ziehe eine zufällige Zahl $r \in [0, 1]$.
 - Falls $r < \frac{p(\vec{x}_i + \vec{d})}{p(\vec{x}_i)}$, akzeptiere $x_{i+1} = \vec{x}_i + \vec{d}$
 - Sonst $x_{i+1} = x_i$
- Die Verteilung der Positionen konvergiert gegen p !

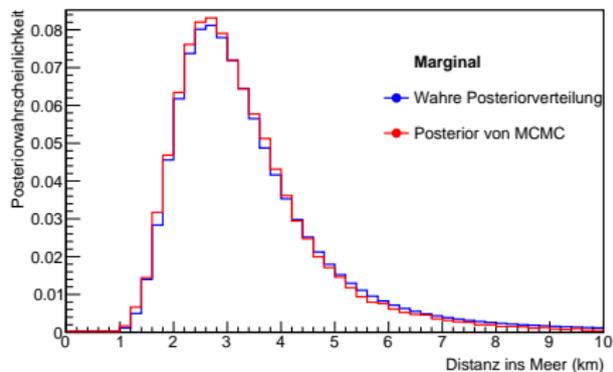


10 Schritte

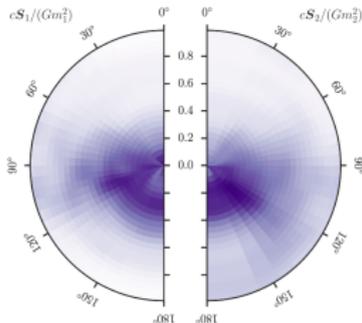
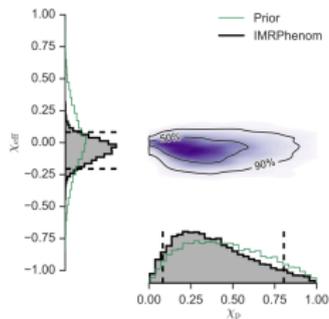
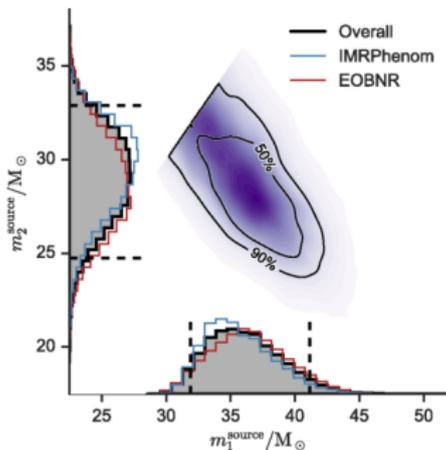
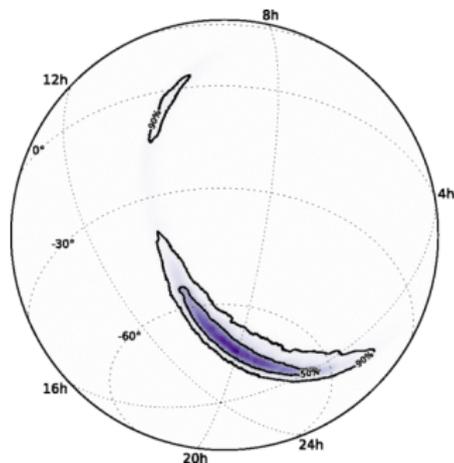
Leuchtturmproblem – Marginale



(tatsächliche Position: (-5.5,3))



Gravitationswellenmessung – Marginale



*Properties of the Binary
Black Hole Merger
GW150914, LIGO
Kollaboration*

- Bayesianische Schlussfolgerung besteht aus einem Schritt: Schreibe Bayes' Theorem hin
- Gibt Wahrscheinlichkeitsverteilung für Gesamtheit an Parametern
- Für Unsicherheit einzelner Parameter: Marginalisierung
- Marginalisierung hochdimensionaler Probleme → Markov-Chain Monte Carlo Methoden
- Vielzahl solcher Methoden, hier nur kurze Einführung

Übersicht über Vorlesung

- Wie kann man Daten sinnvoll darstellen?
- Warum ist es wichtig, Unsicherheiten zu betrachten?
- Was sind statistische und systematische Unsicherheiten?
- Was sind Binomial-, Poisson- und Normalverteilung und wann tauchen sie auf?
- Wie ist Wahrscheinlichkeit definiert?
- Was sind Störparameter?
- Wie charakterisiert man Schätzfunktionen?
- Was gibt der p-Wert an?
- Was ist ein Konfidenzintervall?
- Wie funktioniert Bayesianische Schlussfolgerung?
- Was ist der Prior?
- Was ist ein credible interval?
- Was ist Marginalisierung und wie helfen Markov Chain Monte Carlo Methoden dabei?