

Kapitel 5

Die Maximum-Likelihood-Methode

Die Maximum-Likelihood-Methode wird benützt, um die unbekannt Parameter einer Theoriefunktion aus gemessenen Daten zu bestimmen. Sie ist allgemeiner als die Methode der kleinsten Quadrate (Kap. 4) und auch dort anwendbar, wo die Fehler *nicht* normalverteilt sind. Dies gilt insbesondere bei der Beobachtung seltener Ereignisse, wo die Poisson-Verteilung *nicht* durch eine Gauss-Verteilung angenähert werden kann (Kap. 2.2, 2.2.1). Ein Nachteil der Maximum-Likelihood-Methode im Vergleich zur der Methode der kleinsten Quadrate ist, dass das Maximum der Likelihoodfunktion nichts aussagt über die Güte der Anpassung. Die Maximum-Likelihood-Methode haben wir schon im Kap. 1.3.5 kennengelernt und angewandt auf ein einfaches Problem mit nur einem Parameter. In diesem Kapitel werden wir Theoriefunktionen mit mehreren Parametern betrachten und nach einer theoretische Einführung (Kap. 5.1) praktische Probleme sowohl grafisch (Kap. 5.2 - 5.4) als auch numerisch (Kap. 5.5) lösen.

5.1 Likelihood- und erweiterte Likelihoodfunktion

In einem Experiment werden N Ereignisse

$$\vec{x}_i, \quad i = 1 \cdots N$$

beobachtet im Beobachtungsraum (Phasenraum) Ω . \vec{x}_i sind kinematische Variablen wie z.B. Zeit, Energie und Impulsbetrag: $\vec{x}_i = (t_i, E_i, p_i)$. Das Experiment wird durch eine Theoriefunktion

$$f(\vec{x}; \vec{a}) \quad \text{mit} \quad \int_{\Omega} f(\vec{x}; \vec{a}) d\vec{x} = \mu(\vec{a}) \quad (5.1)$$

beschrieben, wobei $\vec{a} = (a_1, \dots, a_m)$ die unbekannt Parameter sind wie z.B. Masse und Lebensdauer: $\vec{a} = (M, \tau)$. Die Wahrscheinlichkeit N Ereignisse im Phasenraum Ω zu beobachten ist durch die Poisson-Statistik (Kap. 2.2) gegeben

$$P(N; \mu) = \frac{\mu^N e^{-\mu}}{N!} \quad (5.2)$$

Die Wahrscheinlichkeitsdichte im Phasenraum ist

$$p(\vec{x}; \vec{a}) = f(\vec{x}; \vec{a}) / \mu(\vec{a}) \quad (5.3)$$

Die Wahrscheinlichkeit, dass ein Ereignis im Intervall $(\vec{x}, \vec{x} + d\vec{x})$ liegt, ist $p(\vec{x}; \vec{a}) d\vec{x}$. Die Likelihoodfunktion ist proportional zur Gesamtwahrscheinlichkeit der beobachteten Verteilung

$$L = \prod_{i=1}^N p(\vec{x}_i; \vec{a}) \quad (5.4)$$

Das Prinzip der Maximum-Likelihood-Methode ist, dass die beobachtete Verteilung die grösste Wahrscheinlichkeit hat. L ist eine Funktion der Parameter \vec{a} und die besten Schätzwerte für \vec{a} sind diejenigen, die L maximal machen. Statt das Maximum von L kann man auch das Maximum von

$$\boxed{\ln L = \sum_{i=1}^N \ln p(\vec{x}_i; \vec{a})} \quad (5.5)$$

bestimmen (Grund: das Produkt L kann sehr gross werden). Die Schätzwerte \vec{a}_0 der Parameter \vec{a} erhält man somit aus der Lösung der Likelihoodgleichung

$$\left(\frac{\partial \ln L}{\partial a_j} \right)_{a_j = a_{0j}} = 0 \quad \text{für} \quad j = 1 \cdots m \quad (5.6)$$

Die erweiterte Likelihoodfunktion EL ist die Likelihoodfunktion L multipliziert mit der Wahrscheinlichkeit N Ereignisse zu beobachten

$$EL = P(N; \mu) L = \frac{\mu^N e^{-\mu}}{N!} \prod_{i=1}^N p(\vec{x}_i; \vec{a}) \quad (5.7)$$

Auch für die erweiterte Likelihoodfunktion gilt, dass die besten Schätzwerte für \vec{a} diejenigen sind, die $\ln EL$ maximal machen.

$$\begin{aligned} \ln EL &= N \ln \mu - \mu - \ln N! + \sum_{i=1}^N \ln p(\vec{x}_i; \vec{a}) \\ &\propto \sum_{i=1}^N \ln f(\vec{x}_i; \vec{a}) - \mu(\vec{a}) \end{aligned} \quad (5.8)$$

wobei der von den Parametern \vec{a} unabhängige Term $\ln N!$ weggelassen wurde.

Ist μ eine von \vec{a} unabhängige Variable, so ist die erweiterte Likelihoodfunktion EL das Produkt zweier unabhängigen Wahrscheinlichkeiten $P(N, \mu)$ und $L(\vec{a})$. In diesem Fall sind μ und \vec{a} unkorreliert und der Schätzwert für μ ist

$$\left(\frac{\partial \ln EL}{\partial \mu}\right)_{\mu=\mu_0} = \frac{N}{\mu_0} - 1 = 0 \Rightarrow \mu_0 = N \quad (5.9)$$

mit dem Fehler (Gl. 5.19)

$$\sigma_\mu^2 = \frac{-1}{\partial^2 \ln EL / \partial \mu^2} = \mu_0^2 / N = N \quad (5.10)$$

Damit haben wir eine Schätzung für den Mittelwert und die Varianz der Poisson-Verteilung (Kap. 2.2) mit der Likelihoodmethode bestimmt.

Oft sind die N Ereignisse in einem Histogramm mit n Bins

$$\vec{x}_k, \quad y_k, \quad k = 1 \cdots n$$

zusammengefasst, wobei die Binbreite $\Omega_k = \Delta\Omega$ konstant sei. Alle Ereignisse \vec{x}_i mit $\vec{x}_k - \Delta\vec{x}_k/2 < \vec{x}_i < \vec{x}_k + \Delta\vec{x}_k/2$ werden also im Bin (Intervall) $\Omega_k = (\vec{x}_k - \Delta\vec{x}_k/2, \vec{x}_k + \Delta\vec{x}_k/2)$ zusammengefasst. Die y_k Ereignisse in Bin k erhalten dabei alle den gemeinsamen Wert \vec{x}_k . Dabei geht Information verloren. Dieser Verlust ist gering, solange die Binbreite klein ist im Vergleich mit der experimentellen Auflösung. Die erwartete Anzahl der Ereignisse in Bin k ist (Gl. 5.1)

$$\mu_k(\vec{a}) = \int_{\Omega_k} f(\vec{x}; \vec{a}) d\vec{x} \simeq f(\vec{x}_k; \vec{a}) \Delta\Omega \quad (5.11)$$

Summiert über alle Bins erhalten wir

$$\sum_{k=1}^n y_k = N \quad \text{und} \quad \sum_{k=1}^n \mu_k = \mu$$

Die Wahrscheinlichkeit y_k Ereignisse im Bin k zu beobachten ist (Gl. 5.2)

$$P(y_k; \mu_k) = (\mu_k^{y_k} / y_k!) e^{-\mu_k} \quad (5.12)$$

Die erweiterte Likelihoodfunktion ist in diesem Fall

$$EL' = \prod_{k=1}^n P(y_k; \mu_k) = e^{-\mu} \prod_{k=1}^n \mu_k^{y_k} / y_k! \quad (5.13)$$

und der Logarithmus davon (wobei Terme, die unabhängig von \vec{a} sind, weggelassen werden)

$$\begin{aligned} \ln EL' &= \sum_{k=1}^n y_k \ln \mu_k - \mu \\ &\simeq \sum_{k=1}^n y_k \ln f(\vec{x}_k; \vec{a}) - \mu(\vec{a}) \end{aligned} \quad (5.14)$$

Ist das Phasenraumintegral $\mu(\vec{a}) = N_\Omega$ eine Konstante, so erhalten wir den Logarithmus der Likelihoodfunktion für ein Histogramm

$$\boxed{\ln L' = \sum_{k=1}^n y_k \ln p(\vec{x}_k; \vec{a})} \quad (5.15)$$

Gibt es viele Ereignisse pro Bin, so kann in guter Näherung die Poisson-Verteilung durch eine Gauss-Verteilung ersetzt werden (Kap. 2.2)

$$P(y_k; \mu_k) = \frac{1}{\sqrt{2\pi} \sigma} \exp[-(y_k - \mu_k)^2 / 2\sigma^2] \quad \text{mit} \quad \sigma^2 = \mu_k \simeq y_k \quad (5.16)$$

Damit erhalten wir

$$\ln EL' \propto -\frac{1}{2} \sum_{k=1}^n \frac{[y_k - \mu_k(\vec{a})]^2}{y_k} \quad (5.17)$$

In diesem Fall ist die Maximum-Likelihood-Methode äquivalent mit der Methode der kleinsten Quadrate (Kap. 4). Dies ist natürlich auch der Fall, wenn die Wahrscheinlichkeitsverteilung selber schon eine Gauss-Verteilung ist. Hat man z.B. N Messwerte $x_i, y_i \pm \sigma_i$ mit der Verteilung (vgl. Gl. 5.3)

$$p(y(x); \vec{a}) = \frac{1}{\sqrt{2\pi} \sigma_y} \exp(-[y - f'(x, \vec{a})]^2 / 2\sigma_y^2)$$

so ist der Logarithmus der Likelihoodfunktion

$$\ln L \propto -\frac{1}{2} \sum_{i=1}^N \frac{[y_i - f'(x_i, \vec{a})]^2}{\sigma_i^2} = -\frac{1}{2} \chi^2$$

Die Fehler der Parameter bei der Methode der kleinsten Quadrate erhält man aus der Zunahme von χ^2 (siehe Kap. 4.1.1)

$$\chi^2(a_0) = \chi_{min}^2 \quad \boxed{\chi^2(a_0 + \sigma_a) = \chi_{min}^2 + 1} \quad \chi^2(a_0 + 2\sigma_a) = \chi_{min}^2 + 4$$

Entsprechend werden die Fehler der Parameter in der Maximum-Likelihood-Methode definiert

$$\begin{aligned} \ln L(a_0) &= \ln L_{max} & \boxed{\ln L(a_0 + \sigma_a) = \ln L_{max} - 1/2} \\ \ln L(a_0 + 2\sigma_a) &= \ln L_{max} - 2 \end{aligned} \quad (5.18)$$

oder in analytischer Form

$$\frac{\partial^2 \chi^2}{\partial a^2} = \frac{2}{\sigma_a^2} \Rightarrow \boxed{\frac{\partial^2 \ln L}{\partial a^2} = \frac{-1}{\sigma_a^2}} \quad (5.19)$$

Man beachte, dass nur im Falle der Gauss-Verteilung das 1σ -Intervall einer Wahrscheinlichkeit von 68% entspricht.

5.2 Beispiel 1: Exponentialkurve

In einem Experiment werden mit einem TDC N Zerfallszeiten

$$t_i, \quad i = 1 \dots N$$

gemessen, wobei der Phasenraum der Zeitbereich des TDCs ist $\Omega = (0, T)$. Die theoretische Verteilung ist (vgl. Gl. 5.1)

$$f(t; \mu, \lambda) = \mu \frac{\lambda e^{-\lambda t}}{1 - e^{-\lambda T}} \equiv \mu p(t; \lambda) \quad \int_0^T f(t) dt = \mu \quad (5.20)$$

Damit ist das Phasenraumintegral μ unabhängig von λ und sein Schätzwert ist

$$\mu_0 = N \pm \sqrt{N}$$

Im Gegensatz dazu ist bei der "naiven" Theoriefunktion $f'(t; N_0, \lambda) = N_0 \exp(-\lambda t)$ das Phasenraumintegral $\mu' = N_0[1 - \exp(-\lambda T)]/\lambda$ eine Funktion der beiden korrelierten Variablen N_0 und λ (siehe auch die entsprechende Diskussion in Kap. 4.2.1).

Die Zerfallsrate λ erhält man als Lösung der Likelihoodgleichung (Gln. 5.5, 5.6)

$$\begin{aligned} \ln L(\lambda) &= \sum_{i=1}^N \ln p(t_i; \lambda) \\ &= N [\ln \lambda - \lambda \bar{t} - \ln(1 - e^{-\lambda T})] \end{aligned} \quad (5.21)$$

$$\frac{d \ln L}{d \lambda} = N [1/\lambda - \bar{t} - T e^{-\lambda T} / (1 - e^{-\lambda T})] \quad (5.22)$$

wobei $\bar{t} = (1/N) \sum t_i$, $i = 1 \dots N$ der Mittelwert der gemessenen Zeiten ist (Gl. 1.30). Sind die Zeiten in einem Histogramm zusammengefasst (t_k, y_k , $k = 1 \dots n$, $\sum y_k = N$), so ist $\bar{t} = (1/N) \sum y_k t_k$ (Gl. 1.19).

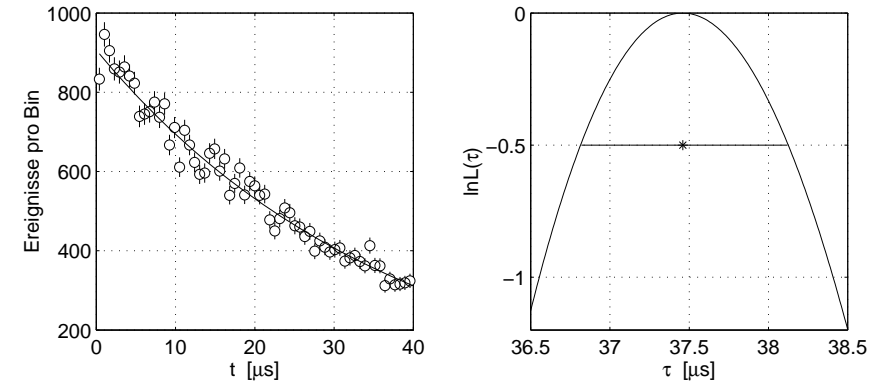


Abbildung 5.1: Zeitspektrum von zufälligen Ereignissen mit Likelihood-Fit

Nur für $T \gg 1/\lambda$ ist eine analytische Lösung möglich

$$\lambda_0 = 1/\tau_0 = 1/\bar{t} \quad \sigma_\lambda^2 = \frac{-1}{d^2 \ln L / d \lambda^2} = \lambda_0^2 / N$$

Das Resultat ist der Mittelwert und Varianz/ N der Exponentialverteilung (Kap. 2.3). In der Praxis ist T von der gleichen Größenordnung wie τ , da man für $T \gg \tau$ nur noch Untergründereignisse misst. Ein Schätzwert für λ erhält man am einfachsten durch die graphische Darstellung von $\ln L(\lambda)$. Der λ -Wert des Maximums und dessen Fehler (Gl. 5.18) kann man dann ablesen. Ein Beispiel ist in Abb. 5.1 (Programm 5.1) gegeben. Das Resultat $\tau_0 = 1/\lambda_0 = (37.4 \pm 0.7) \mu\text{s}$ ist in Übereinstimmung mit dem Resultat aus dem χ^2 -Fit an den gleichen Daten (Kap. 4.2.1).

```

%-----
%      program 5.1: likelihood analysis of exponential distribution
%      -----
subplot(121),
load exp_data1,
hold on,
plot_data(exp_data1)
grid

```

```

axis('square'),
xlabel('t [\mus]'),
x = exp_data1(:,1);
dt = x(2) - x(1)
n = length(x);
N = sum(y)
out = centroid(x,y)
tau = 36.5:0.01:38.5;
LH = N*(-log(tau) - out(1) ./tau - log(1 - exp(-T ./ tau)));
[m,i] = max(LH);
norm = sum(y)*dt/(1 - exp(-T/tau0));
plot(x,(norm/tau0)*exp(-x/tau0),'b-'),
plot(tau,LH-m,'r-'),
xlabel('\tau [\mus]'),
axis('square'),
hold on,
c = LH-m >= -0.5;
%-----
axis([0 40 200 1000])
ylabel('Ereignisse pro Bin')
y = exp_data1(:,2);
% 0.632 microsec/channel
T = x(n) + dt/2
% number of events
% out(1) = mean value
tau0 = tau(i)
subplot(122)
grid
ylabel('lnL(\tau)')
plot(tau0,-0.5,'*')
plot(tau(c),0*tau(c)-0.5,'-')

```

5.3 Beispiel 2: Signal und Untergrund

In einem Experiment wird ein Signal über einem Untergrund gesucht. Die Form des Signals $f_1(\vec{x})$ und des Untergrunds $f_2(\vec{x})$ werden als bekannt vorausgesetzt. Damit gibt es zwei Parameter: N_1 die gesuchte Zahl der Signal-Ereignisse und N_2 die Zahl der Untergrund-Ereignisse. Die Theoriefunktion ist damit (Gl. 5.1))

$$f(\vec{x}; N_1, N_2) = N_1 f_1(\vec{x}) + N_2 f_2(\vec{x}) \quad (5.23)$$

Die Normierung ist

$$\int_{\Omega} f_j(\vec{x}) d\vec{x} = 1 \quad (j = 1, 2) \quad \text{und damit} \quad \int_{\Omega} f(\vec{x}; N_1, N_2) d\vec{x} = N_1 + N_2 \quad (5.24)$$

Die Wahrscheinlichkeitsdichte ist

$$p(\vec{x}; N_1, N_2) = f(\vec{x}; N_1, N_2)/(N_1 + N_2) \quad (5.25)$$

Die Likelihoodfunktion $\ln L$ (Gl. 5.5) beschreibt die Verteilung der N gemessenen Ereignisse im Phasenraum Ω , wobei $N = N_1 + N_2$ konstant ist. Sie hängt also nur von einem Parameter N_1 ab ($N_2 = N - N_1$)

$$\ln L(N_1) = \sum_{i=1}^N \ln p(\vec{x}_i; N_1) \propto \sum_{i=1}^N \ln [N_1 f_1(\vec{x}_i) + (N - N_1) f_2(\vec{x}_i)] \quad (5.26)$$

Die erweiterte Likelihoodfunktion $EL = P(N; N_1 + N_2) L$ (Gl. 5.7) hängt von beiden Variablen ab (Gl. 5.8)

$$\ln EL(N_1, N_2) \propto \sum_{i=1}^N \ln f(\vec{x}_i; N_1, N_2) - (N_1 + N_2) \quad (5.27)$$

Für histogrammierte Ereignisse ($\sum_{i=1}^n y_i = N$) wird $\sum_{i=1}^N \ln f(x_i)$ durch $\sum_{i=1}^n y_i \ln f(x_i)$ ersetzt (Gl. 5.15 bzw. 5.14).

Es ist klar, dass die Poisson-Verteilung $P(N; N_1 + N_2)$ nicht zwischen N_1 und N_2 unterscheiden kann. Die Likelihoodfunktion und die erweiterte Likelihoodfunktion geben damit die gleichen Schätzwerte für N_1 und N_2 , aber nicht die gleichen Fehler, wie wir am folgenden Beispiel sehen werden.

Gauss-Peak mit exponentiellem Untergrund

Abb. 5.2 zeigt ein Monte-Carlo-generiertes Spektrum bestehend aus einem exponentiellen Untergrund und einem Gauss-Peak. Die Theoriefunktion ist also (Gl. 5.23)

$$f(x; N_g, N_e) = \frac{N_g}{\sqrt{2\pi}\sigma} e^{-(x-x_0)^2/2\sigma^2} + N_e \lambda e^{-\lambda x} \quad (5.28)$$

wobei $x_0 = 6$, $\sigma = 0.7$ und $\tau = 1/\lambda = 2$ fest vorgegeben sind. Sowohl eine Likelihood-Analyse (Gl. 5.26) als auch eine Analyse mit der erweiterten Likelihoodfunktion (Gl. 5.27) wurde durchgeführt (Abb. 5.2, Programm 5.2). Das Resultat der Likelihood-Analyse ist (Abb. 5.2: Schnittpunkte der 1σ -Kontur mit der Geraden $N_g + N_e = N = 746$; Output Programm 5.2)

$$N_g = 246.5 \pm 14.6 \quad \text{und damit} \quad N_e = N - N_g = 499.5 \pm 14.6 \quad (5.29)$$

Die Analyse mit der erweiterten Likelihoodfunktion ergibt (Abb. 5.2: Tangenten an die 1σ -Kontur; Output Programm 5.2)

$$N_g = 247.0 \pm 17.1 \quad \text{und} \quad N_e = 500.0 \pm 23.5 \quad (5.30)$$

Die Schätzwerte für N_g und N_e sind bei beiden Methoden gleich, ihre Fehler jedoch nicht (der Fehler in Gl. 5.29 ist kleiner als der Poisson-Fehler $\sqrt{N_g}$ und $\sqrt{N_e}$). Die Likelihoodfunktion (Gl. 5.26) ist genau genommen keine Funktion von N_g aber von $r \equiv N_g/N$. Statt Gl. 5.29 ist das Resultat also

$$r \equiv N_g/N = 0.3304 \pm 0.0196 \quad (5.31)$$

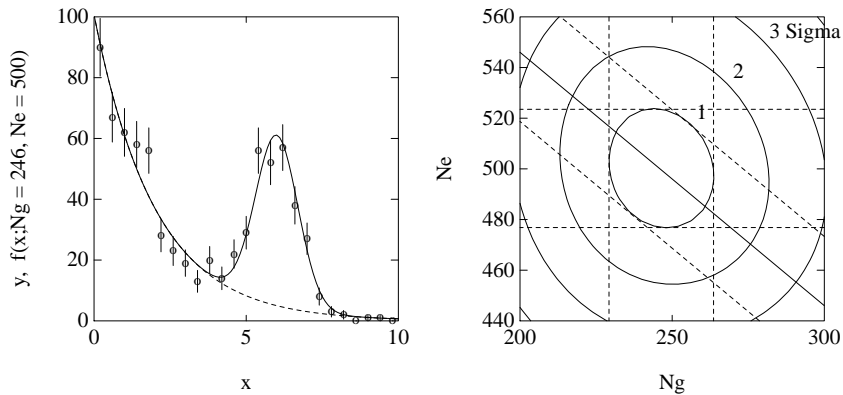


Abbildung 5.2: Links: Untergrundspektrum mit einem Gauss-Peak (Monte-Carlo) und Fitfunktion. Rechts: Erweiterte Likelihoodfunktion als Funktion der Fitparameter. Die 1σ -, 2σ - und 3σ -Konturen und die Linie $N_e + N_g = N$ sind dargestellt. Die Tangenten (- -) an die 1σ -Kontur entsprechen den Fehlern in N_g , N_e und N .

Der Fehler ist etwas grösser als der Binomial-Fehler (Kap. 2.1) $\sigma = \sqrt{r(1-r)/N} = 0.0172$. Man hätte eine Binomialverteilung, wenn der Untergrund und das Signal vollständig getrennt wären. Die Fehler in $N_g = Nr$ und $N_e = N(1-r)$ folgen aus dem Fehlerfortpflanzungsgesetz

$$\sigma_{N_g} = N_g \sqrt{\sigma_r^2/r^2 + \sigma_N^2/N^2} = 17.2$$

und

$$\sigma_{N_e} = N_e \sqrt{\sigma_r^2/(1-r)^2 + 1/N} = 23.4$$

in Übereinstimmung mit der erweiterten Likelihood-Analyse (GL. 5.30).

```

%-----
%   MATLAB program 5.2: 2 parameter extended Likelihood fit: ng, ne
%   -----
load data5,
sig = 0.7;
x = data5(:,1);
dx = x(2) - x(1);
ex = (dx/tau)*exp(-x/tau);
gx = dx*gauss(x,x0,sig);
nga = 200;

tau = 2;
x0 = 6;
y = data5(:,2);
N = sum(y);
% normalized exponential
% normalized gaussian
ngb = 300;

```

```

nea = 440;
ng = nga:4:ngb;
[X,Y] = meshgrid(ng,ne);
for i = 1:length(x)
    fx = (X*gx(i) + Y*ex(i));
    lnEL = lnEL + y(i)*log(fx);
end
lnEL = lnEL - X - Y;
contour(ng,ne,lnEL-max(max(lnEL)),v),
xlabel('Ng'),
hold on
plot([nga ngb],[N-nga N-ngb],'k-')
plot([nga ngb],[N+sqrt(N)-nga N+sqrt(N)-ngb],'k--')
plot([nga ngb],[N-sqrt(N)-nga N-sqrt(N)-ngb],'k--')
disp('specify 1 sigma points (2) on N = constant fit')
[x,y] = ginput(2);
fprintf(' x1 = %6.2f, x2 = %6.2f, dx = %6.2f\n',x(1),x(2),x(2) - x(1))
fprintf(' y1 = %6.2f, y2 = %6.2f, dy = %6.2f\n',y(1),y(2),y(2) - y(1))
disp('specify 1 sigma points (4: max,min,x,y)')
[x,y] = ginput(4);
ya = min(y);
xa = min(x);
plot([nga ngb],[ya ya],'k--'),
plot([xa xa],[nea neb],'k--'),
fprintf(' x1 = %6.2f, xr = %6.2f, dx = %6.2f\n',xa,xb,xb - xa)
fprintf(' yu = %6.2f, yd = %6.2f, dy = %6.2f\n',ya,yb,yb - ya)
%-----
%   output: specify 1 sigma points (2) on N = constant fit
%           x1 = 231.90, x2 = 261.00, dx = 29.10
%           y1 = 514.02, y2 = 484.83, dy = 29.19
%           specify 1 sigma points (4: max,min,x,y)
%           x1 = 229.87, xr = 264.05, dx = 34.17
%           yu = 476.49, yd = 523.41, dy = 46.92
%-----

```

5.4 *Beispiel 3: Untergrund ohne Signal

Die Suche nach "neuer" Physik, wie z.B. der Zerfall des Protons, die Verletzung der Leptonenzahlerhaltung und die Masse des Neutrinos, ist wichtig um den Gültigkeitsbereich des Standardmodells der Teilchenphysik zu bestimmen. Der gesuchte Prozess erscheint dabei oft als Peak (Signal) über einem kontinuierlichen

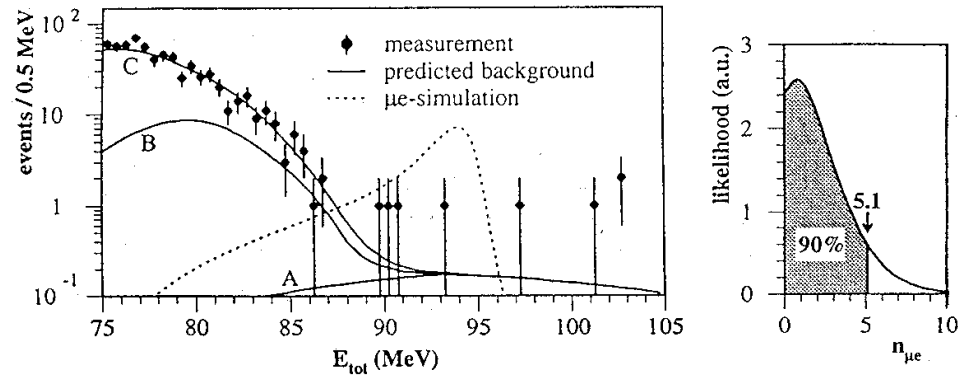


Abbildung 5.3: Links: Gemessenes Energiespektrum der Elektronen und simulierte μe -Konversionslinie (50 Ereignisse). Rechts: Maximum-Likelihood-Analyse [Hone 96].

Untergrund. Als Beispiele werden wir zwei Experimente mit den SINDRUM I und SINDRUM II Detektoren am PSI besprechen.

$\mu \rightarrow e$ Konversion in Blei

Mit dem SINDRUM II Detektor wurde am PSI nach der neutrinolosen Konversion



gesucht. Dabei entsteht ein monoenergetisches Elektron ($E_e \simeq 95$ MeV). Der Untergrund wird verursacht durch erlaubte μ - und π -Prozesse und durch die kosmische Strahlung. Aus Eichmessungen und Monte-Carlo-Simulation ist die Auflösung des Detektors bekannt und damit die zu erwartende Linienform. Das gemessene Elektronenspektrum (Abb. 5.3) ist mit dem erwarteten Untergrund verträglich und eine Maximum-Likelihood-Analyse ergab eine obere Limite von 5.1 μe -Ereignisse (Abb. 5.3). Insgesamt wurden 1.1×10^{11} μ -Einfänge in Pb untersucht und das Verzweungsverhältnis zum erlaubten Einfangsprozess ist [Hone 96]

$$B_{\mu e} = \frac{N(\mu^- + \text{Pb} \rightarrow e^- + \text{Pb})}{N(\mu^- + \text{Pb} \rightarrow \nu_\mu + \text{TI}^*)} \leq 4.6 \times 10^{-11}$$

Statt diesem Beispiel, das drei Untergrundkomponenten hat, wollen wir ein ähnliches, aber einfacheres, Monte-Carlo-Beispiel behandeln. Im Programm

5.3 werden $N = 200$ exponentialverteilte Untergrundereignisse gewürfelt und anschliessend analysiert. In der Maximum-Likelihood-Analyse wird, wie im Kap. 5.3, ein Gauss-Peak über einem exponentiellen Untergrund gesucht. Die Likelihoodfunktion¹ ist also (Gl. 5.26)

$$\ln L(N_g) = \sum_{i=1}^N \ln [N_g g(x_i) + (N - N_g) e(x_i)] \quad (5.32)$$

Die normierten Funktionen $g(x)$ und $e(x)$ sind (Gl. 5.28)

$$g(x) = \frac{1}{\sqrt{2\pi}\sigma} e^{-(x-x_0)^2/2\sigma^2} \quad e(x) = \lambda e^{-\lambda x} \quad (5.33)$$

wobei $x_0 = 6$, $\sigma = 0.7$ und $\tau = 1/\lambda = 2$ fest vorgegeben sind. Das Maximum der Likelihoodfunktion (Abb. 5.4) kann, abhängig von den statistischen Schwankungen des Untergrunds, sowohl im Bereich $N_g > 0$ als auch im Bereich $N_g < 0$ liegen. Physikalisch erlaubt ist nur der Bereich $N_g > 0$ und mit der Bayes-Methode (siehe Kap. 2.2.1) kann für ein bestimmtes Konfidenzniveau CL eine obere Limite definiert werden

$$\text{CL} = \frac{\int_0^{N_g^{\text{max}}} L(N'_g) dN'_g}{\int_0^\infty L(N'_g) dN'_g}$$

Das Resultat ist $N_g < N_g^{\text{max}} = 9.5$ (90 % CL).

```

%-----
%      MATLAB program 5.3: likelihood fit of gaussian and exponential
%      -----
tau = 2;                                sig = 0.7;
x0 = 6;                                  N = 200;
dx = tau/5;                               xmax = 5*tau;
subplot(121),
xevents = -tau*log(rand(N,1));
x = dx/2:dx:xmax+dx/2;                   y = hist(xevents,x);
plot_data([x' y'], 'g'),                  hold on
axis('square'),                           axis([0 xmax 0 40])
xpl = 0:0.1:xmax;                          fx = N*dx*exp(-xpl/tau)/tau;
plot(xpl, N*dx*gauss(xpl,x0,sig)/3, 'b--'), plot(xpl,fx, 'r-')
xlabel('x'),                               ylabel('dN/dx')

```

¹ Da die Zahl der Untergrundereignisse gross ist im Vergleich zu den zu erwartenden Signalereignissen ergeben Likelihood- und erweiterte Likelihoodfunktion in diesem Fall auch die gleichen Fehlergrenzen.

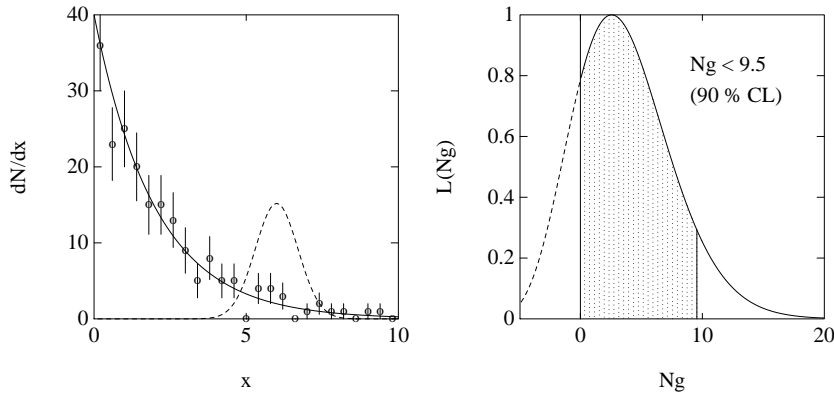


Abbildung 5.4: Untergrundspektrum (Monte-Carlo) und obere Grenze für Anzahl Ereignisse im Gauss-Peak

```

gx = gauss(xevents,x0,sig);
a = -5+0.05;
ng = a:0.1:b;
for i = 1:length(ng)
    fx = ng(i)*gx + (N-ng(i))*ex;
end
L = exp(lnL - max(lnL));
subplot(122),
plot(ng(ng > 0),L(ng > 0),'r-'),
axis('square'),
plot(ng(ng < 0),L(ng < 0),'r--'),
xlabel('Ng'),
ylabel('L(Ng)'),
L = L(ng > 0);
y = zeros(size(ng));
for i=1:length(ng),
end;
c = y/y(1) > 0.1;
for i = 4:4:n,
end;
txt=sprintf('Ng < %3.1f',ng(n)-0.05);
text(9,0.8,txt);
%-----
ex = (1/tau)*exp(-xevents/tau);
b = 20;
lnL = zeros(size(ng));
lnL(i) = sum(log(fx));

hold on
axis([a b 0 1])
plot([0 0],[0 1],'k-')
ylabel('L(Ng)')
ng = ng(ng > 0);

y(i) = sum(L(i:length(ng)));

n = sum(c)+1;
plot([ng(i) ng(i)],[0 L(i)'],'k:')
plot([ng(n) ng(n)],[0 L(n)'],'k-')

text(9,0.7,'(90 % CL)')

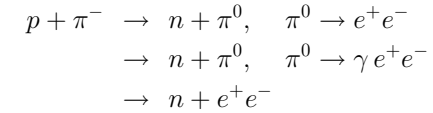
```

Man beachte, dass im Programm 5.3 aus rechentechnischen Gründen (Genauigkeit und Geschwindigkeit) zuerst der Logarithmus $\ln L$ berechnet wird und

anschliessend $L = \exp(\ln L)$.

Seltene π -Zerfälle

Mit dem SINDRUM I Detektor wurde am PSI, neben seltenen und im Standardmodell verbotenen μ -Zerfällen, auch seltene π -Zerfälle untersucht, z.B. folgende Prozesse



Das Hauptinteresse war die Nachweis des Zerfalls $\pi^0 \rightarrow e^+e^-$. Die negativen Pionen eines PSI-Strahls wurden in einem Moderator abgebremst und in einem Flüssig-Wasserstoff-Target gestoppt. Mit Schnitten im Phasenraum wurden e^+e^- -Paare selektiert und gleichzeitig der Untergrund der Dalitz-Zerfälle $\pi^0 \rightarrow \gamma e^+e^-$ stark reduziert. Die Geschwindigkeit der Neutronen $\beta = v/c$ wurde aus einer Flugzeitmessung bestimmt. Für den Zweikörperzerfall des pionischen Wasserstoffatoms $p + \pi^- \rightarrow n + \pi^0$ ist $\beta = 0.0299$. Der Dreikörperzerfall $p + \pi^- \rightarrow n + e^+e^-$ bildet einen kontinuierlichen Untergrund im Geschwindigkeitspektrum. Im gemessenen Spektrum

$$\beta_i, y_i, \quad i = 1 \dots n, \quad \sum_{i=1}^n y_i = N = 565$$

ist kein π^0 -Signal zu sehen (Abb. 5.5) und eine obere Limite für die Zahl der π^0 -Ereignisse wurde mit der Maximum-Likelihood-Methode bestimmt.

Die Form des Signals $g_i = g(\beta_i) \Delta\beta$ wurde mit dem Dalitz-Zerfall gemessen ($\sum_{i=1}^n g_i = 1$). Der Untergrund wird durch eine Gerade beschrieben

$$f_i(\beta_0) = f(\beta_i, \beta_0) \Delta\beta = \frac{\beta_i - \beta_0}{\sum_{i=1}^n \beta_i - n\beta_0} \quad \sum_{i=1}^n f_i = 1$$

wobei β_0 ein Fitparameter ist. Die Likelihoodfunktion (Gl. 5.26) hängt also von zwei Parametern ab

$$\ln L(N_\pi, \beta_0) = \sum_{i=1}^n y_i \ln[N_\pi g(x_i) + (N - N_\pi) f_i(\beta_0)]$$

Der Konturplot (Abb. 5.5) zeigt das Maximum bei $(N_\pi = -6.0, \beta_0 = 1.6)$. Von den beiden Parameter interessiert nur die Zahl der Pionen N_π , deshalb wird über β_0

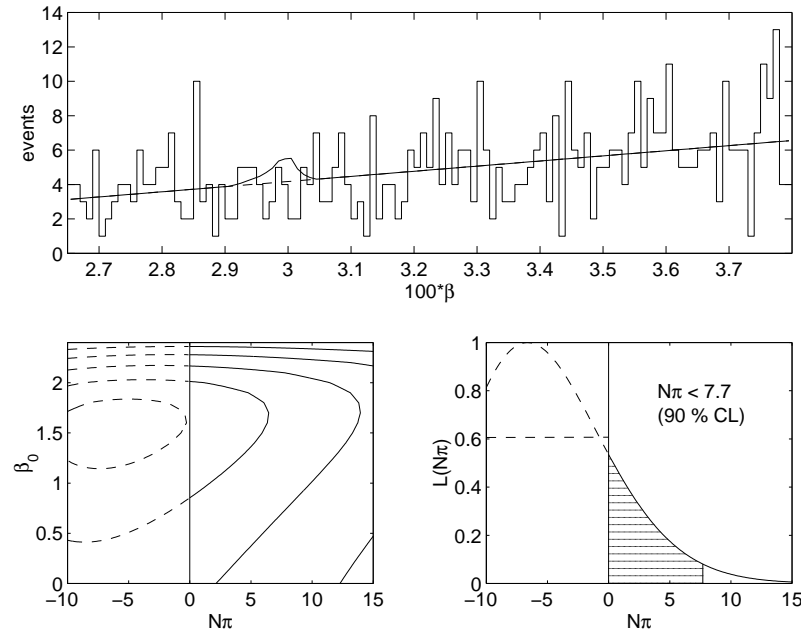


Abbildung 5.5: Suche nach dem Zerfall $\pi^0 \rightarrow e^+e^-$ [Nieb 89]. Oben: Gemessenes Spektrum mit der Fitfunktion (557 Untergrund- und 8 Signalereignisse). Unten: zwei- und eindimensionale Likelihoodfunktionen mit $k\sigma$ -Konturen, bzw. obere Limite für die Zahl der Signalereignisse.

integriert (im Programm 5.4 durch Summation)

$$L(N_\pi) = \int L(N_\pi, \beta_0) d\beta_0.$$

Das Maximum ($N_\pi = -6.0$) dieser Funktion liegt ein σ im negativen nicht-physikalischen Bereich. Das 1σ -Konfidenzintervall folgt aus Gl. 5.18

$$\ln L(a_0 + \sigma_a) = \ln L(a_0) - 1/2 \quad \Rightarrow \quad \frac{L(a_0 + \sigma_a)}{L(a_0)} = 0.607 \quad (5.34)$$

Die obere Limite wird aus dem physikalisch erlaubten Teil des Likelihoodfunktions bestimmt. Das Resultat $N_\pi < 7.7$ (Cl = 90 %) ist in Übereinstimmung mit dem theoretischen Erwartungswert für die 6×10^7 untersuchten π_0 -Zerfälle, $N_{th} = 4$.

```

%-----
%   MATLAB program 5.4: likelihood fit of pi0 experiment
%   -----
subplot(211),
load pi0dataxy
xd = pi0dataxy(:,1);
nd = length(xd);
stairs([xd-0.005;3.8],[yd;y(nd)]),
load pi0peak
xp = pi0peak(:,1);
Left = xd < xp(1);
g = [zeros(sum(Left),1); yp; zeros(sum(Right),1)];
xlabel('100*\beta'),
b0 = 0:0.1:2.4;
[X,Y] = meshgrid(Np,b0);
for i = 1:nd
    fi = (xd(i) - Y)./(sum(xd) - nd*Y);
    lnL = lnL + yd(i)*log(X*g(i) + (N-X).*fi);
end
subplot(223)
v = -0.5*(1:5).^2;
lnLneg = lnL(:,Np < 0);
contour(Np(Np < 0),b0,lnLneg,v,'--'),
axis([-10 15 0 2.4])
contour(Np(Np >= 0),b0,lnLpos,v,'-'),
xlabel('N\pi'),
subplot(224),
L2 = exp(lnL);
plot(Np(Np < 0),L(Np < 0),'--'),
axis([-10 15 0 1])
L = L(Np >= 0);
plot(Np,L,'-'),
plot([-10 0],[0.606 0.607],'k--')
xlabel('N\pi'),
y = zeros(size(Np));
for i=1:length(Np),
end;
c = y/y(1) > 0.1;
for i = 1:n,
end;
txt=sprintf(' < %3.1f',Np(n));
text(4,0.7,'(90 % CL)'),
[Np(n-3:n+3)' y(n-3:n+3)']/y(1)

yd = pi0dataxy(:,2);
N = sum(yd);
axis([2.65 3.8 0 14])
yp = pi0peak(:,2);
Right = xd > xp(length(xp));
ylabel('events')
Np = -10:0.1:15;
lnL = zeros(size(X));
lnL = lnL-max(max(lnL));
lnLpos = lnL(:,Np >= 0);
hold on,
plot([0 0],[0 2.4],'k-')
ylabel('\beta_0')
L = sum(L2)/max(sum(L2));
hold on
Np = Np(Np >= 0);
plot([0 0],[0 1],'k-')
ylabel('L(N\pi)')
y(i) = sum(L(i:length(Np)));
n = sum(c)+1;
plot([Np(i) Np(i)],[0 L(i)'],'k:')
plot([Np(n) Np(n)],[0 L(n)'],'k-')
text(4,0.8,['N\pi' txt])

```



```
subplot(211), hold on, b0 = 1.6;
f = (N-Np(n))*(xd-b0)/(sum(xd)-nd*b0); g = Np(n)*g;
plot(xd,f+g,'r-'), plot(xd,f,'r--')
axis([2.65 3.8 0 14])
%-----
```

5.5 *Numerische Lösungen

Die gleichen numerischen Methoden, die iterativ das Minimum einer χ^2 -Funktion bestimmen, können auch benützt werden um das Maximum einer Likelihoodfunktion zu bestimmen.

Linearisierung der Likelihoodfunktion (vgl. Kap. 4.2.2)

In der Nähe des Maximums $\vec{a} = \vec{a}_0 + \delta\vec{a}$ kann man die Likelihoodfunktion (Gl. 5.5 oder 5.15) linearisieren durch eine Taylor-Entwicklung bis zur ersten Ordnung

$$\ln L = \sum_{i=1}^n \left[y_i \ln p_0(x_i) + \sum_{k=1}^m \frac{y_i}{p_0(x_i)} \frac{\partial p_0(x_i)}{\partial a_k} \delta a_k \right], \quad p_0(x_i) \equiv p(x_i; \vec{a}_0) \quad (5.35)$$

Die m gekoppelten Likelihoodgleichungen werden damit

$$\frac{\partial \ln L}{\partial a_j} = \sum_{i=1}^n \frac{y_i}{p_0(x_i)} \frac{\partial p_0(x_i)}{\partial a_j} - \sum_{i=1}^n \sum_{k=1}^m \frac{y_i}{p_0^2(x_i)} \frac{\partial p_0(x_i)}{\partial a_j} \frac{\partial p_0(x_i)}{\partial a_k} \delta a_k = 0 \quad (5.36)$$

Die Lösung erfolgt wieder durch Matrizen-Inversion

$$\boxed{\delta\vec{a} = S^{-1}\vec{b} = C\vec{b}} \quad (5.37)$$

Die Komponenten des m -dimensionalen Vektors \vec{b} und der $m \times m$ -Matrix S sind (Gl. 5.36)

$$\boxed{b_j = \sum_{i=1}^n \frac{y_i}{p_0(x_i)} \frac{\partial p_0(x_i)}{\partial a_j}} \quad \boxed{S_{jk} = \sum_{i=1}^n \frac{y_i}{p_0^2(x_i)} \frac{\partial p_0(x_i)}{\partial a_j} \frac{\partial p_0(x_i)}{\partial a_k}} \quad (5.38)$$

Diese Gleichungen (vgl. Gl. 4.31) sind in der MATLAB-Funktion `mrqmax` (vgl. MATLAB-Funktion `mrqmin` in Programm 4.7) programmiert. Der Marquardt-Algorithmus, der diese Funktion aufruft, ist in der MATLAB-Funktion `LHmaximum`

(vgl. `LSminimum`) programmiert.

Simplex-Methode

Wenn man das Vorzeichen der Likelihoodfunktion umdreht, kann man die Simplex-Methode, die für die χ^2 -Funktion benützt wurde (Kap. 4.2.5), direkt einsetzen. Ein Beispiel, das früher graphisch gelöst wurde (Programm 5.2, Abb. 5.2), ist in Programm 5.5 gegeben. Die zu optimierende Funktion (Gl. 5.27, modifiziert für histogrammierte Ereignisse) ist `lnEL = SBlikelihood(par)` und die Theoriefunktion (Gl. 5.28) `fx = SBfunction(x,p)`. Nach dem Optimierung mit dem Simplex-Methode `a = fmins('SBlikelihood',a,[1 0.1 0.1])` werden die Fehler der Parameter mit Hilfe der Gl. 5.38 berechnet. Dazu werden in der Funktion `fx = SBfunction(x,p)` auch die Ableitungen $\partial f(x)/\partial N_g$ und $\partial f(x)/\partial N_e$ berechnet (zweite bzw. dritte Spalte der Matrix `fx = [Ng*gx+Ne*ex gx ex]`). Das Resultat $N_g = 246 \pm 17$ und $N_e = 500 \pm 23$ ist in Übereinstimmung mit dem Resultat der grafischen Analyse (Kap. 5.3). Zusätzlich wird auch der Korrelationskoeffizient $\rho = -12\%$ berechnet. Man beachte, dass die erweiterte Likelihoodfunktion (Gl. 5.14) benützt wurde (mit negativem Vorzeichen: `lnEL = -sum(y .* log(fx1)) + sum(par)`). Die Linearisierung dieser Funktion führt zu der gleichen S -Matrix (Gl. 5.38) mit $p_0(x_i) \equiv p(x_i; \vec{a}_0)$ (Gl. 5.3) ersetzt durch $f_0(x_i) \equiv f(x_i; \vec{a}_0)$ (Gl. 5.1).

```
%-----
% program 5.5: likelihood fit of signal and background
% -----
load data5;
Gdata = data5;
global Gdata
a = [200 550]';
a = fmins('SBlikelihood',a,[1 0.1 0.1]);
x = Gdata(:,1); y = Gdata(:,2);
fx = SBfunction(x,a); g = y ./ fx(:,1).^2;
S = zeros(length(a),length(a));
for j = 1:length(a)
    for k = j:length(a)
        S(j,k) = sum(g.*fx(:,k+1).*fx(:,j+1)); S(k,j) = S(j,k);
    end
end;
C = inv(S);
fprintf('\nNg = %5.1f +-%5.1f, ',a(1),sqrt(C(1,1)))
fprintf('\nNe = %5.1f +-%5.1f, ',a(2),sqrt(C(2,2)))
fprintf('\rho = %5.3f\n',C(1,2)/sqrt(C(2,2))/sqrt(C(1,1)))
```

```

%-----
function lnEL = SBlikelihood(par)
%-----
global Gdata
x = Gdata(:,1);          y = Gdata(:,2);
fx = SBfunction(x,par);  fx1 = fx(:,1);
lnEL = -sum(y .* log(fx1)) + sum(par);
%-----

function fx = SBfunction(x,p)
%-----
[n m] = size(x);
if n < m,                x = x';
end;                      tau = 2;
sig = 0.7;               x0 = 6;
Ng = p(1);               Ne = p(2);
dx = x(2) - x(1);        ex = (dx/tau)*exp(-x/tau);
gx = dx*gauss(x,x0,sig);  fx = [Ng*gx+Ne*ex gx ex];
%-----

%   Ng = 246.4 +- 17.2,   Ne = 499.5 +- 23.4,   rho = -0.121
%-----

```

Ein komplizierteres Beispiel, das schon mit der Methode der kleinsten Quadrate berechnet wurde (Marquardt-Methode, Kap. 4.2.4 und Simplex-Methode, Kap. 4.2.5) ist in Programm 5.6 gegeben. Die Theoriefunktion $fx = EGfunc2(x,p)$ ist richtig normiert für die endliche Messzeit $T = 12$ (Faktor $1/(1 - \exp(-\lambda*(12 - x_0)))$) und die Ableitungen werden numerisch berechnet $dfda(:,j) = (EGfunc2(x,a+da) - EGfunc2(x,a-da))/0.002$. Für die Likelihoodfunktion $\ln L$ ist es notwendig, dass die Theoriefunktion, integriert über den Phasenraum des Experiments (also nicht bis ∞), auf Eins normiert ist, unabhängig von den Parametern. Für die χ^2 -Funktion ist dies nicht erforderlich und deshalb konnte eine einfachere Theoriefunktion und damit analytische Ableitungen benutzt werden (siehe Funktion $fx = EGfunction(x,p)$ in Programm 4.9). Das Resultat der Likelihoodanalyse (Programm 5.6) ist das gleiche wie das Resultat der χ^2 -Analyse (Programm 4.9 und 4.10).

```

%-----
%   MATLAB program 5.6: Likelihood fit of exponential folded with
%   gaussian using MATLAB function fmins (simplex algorithm)
%   -----
format compact
load EGdata;
x = EGdata(:,1);          y = EGdata(:,2);
global EGdata
logplot_data(EGdata),    hold on
axis([0 12 1 1000])
t = 0:0.2:40;             a = [1 1 3]';
fx = EGfunc2(t,a);        semilogy(t,fx(:,1),'g--');
a = fmins('EGlikelihood2',a,[1 0.01 0.1]);
fx = EGfunc2(t,a);        semilogy(t,fx,'b-');
fx = EGfunc2(x,a);        g = y ./ fx.^2;
dfda = zeros(length(x),length(a));  da = zeros(size(a));
for j = 1:length(a)
    da(j) = 0.001;
    dfda(:,j) = (EGfunc2(x,a+da) - EGfunc2(x,a-da))/0.002;
    da(j) = 0;
end
S = zeros(length(a),length(a));
for j = 1:length(a)
    for k = j:length(a)
        S(j,k) = sum(g.*dfda(:,k).*dfda(:,j)); S(k,j) = S(j,k);
    end
end;
C = inv(S);
fprintf('\nlambda = %5.3f +- %5.3f, ',a(1),sqrt(C(1,1)))
fprintf('   sigma = %5.3f +- %5.3f, ',a(2),sqrt(C(2,2)))
fprintf('   x0 = %5.3f +- %5.3f\n',a(3),sqrt(C(3,3)))
%-----
%   lambda = 0.499 +- 0.010,   sigma = 0.708 +- 0.021,   x0 = 2.007 +- 0.027
%-----

function lnL = EGLikelihood2(par)
%-----
global EGdata
x = EGdata(:,1);          y = EGdata(:,2);
fx = EGfunc2(x,par);      lnL = -sum(y .* log(fx));
%-----

```

```
function fx = EGfunc2(x,p)
%-----
[n m] = size(x);
if n < m,          x = x';          end
lambda = p(1);    sigma = p(2);    x0 = p(3);
norm = 1191;
u = (x - x0)/(sqrt(2)*sigma) - lambda*sigma/sqrt(2);
fx = 0.5*lambda*exp(-lambda*(x - x0 - 0.5*lambda*sigma^2));
fx = norm*(1 + erf(u)).*fx / (1 - exp(-lambda*(12 - x0)));
%-----
```

5.6 *Magische Formel

Mit der Likelihoodfunktion kann man die Fehler der Parameter berechnen (Gl. 5.19). Dies ist sogar a priori möglich, d.h. bevor man experimentelle Daten hat. Neben einer aufwendigen Monte-Carlo-Simulation ist dies eine sehr wichtige Möglichkeit die Genauigkeit eines geplanten Experiments abzuschätzen.

Für N gemessene Ereignisse ist (Gl. 5.5 in Gl. 5.19 eingesetzt)

$$\frac{\partial^2 \ln L}{\partial a_j \partial a_k} = \sum_{i=1}^N \frac{1}{p(x_i)} \frac{\partial^2 p(x_i)}{\partial a_j \partial a_k} - \sum_{i=1}^N \frac{1}{p^2(x_i)} \frac{\partial p(x_i)}{\partial a_j} \frac{\partial p(x_i)}{\partial a_k} \quad (5.39)$$

Für $N \rightarrow \infty$ geht eine gemessene Verteilung x_i über in eine theoretische Verteilung $p(x)$ und der gemessene Mittelwert einer Funktion $f(x_i)$ (vgl. Gl. 1.30) geht über in der Erwartungswert von $f(x)$ (vgl. Gl. 1.11)

$$\frac{1}{N} \sum_{i=1}^N f(x_i) \simeq E[f(x)] = \int f(x) p(x) dx$$

In dieser Näherung ($N \geq 100$ statt $N = \infty$) verschwindet der erste Term der Gl. 5.39

$$\sum_{i=1}^N \frac{1}{p(x_i)} \frac{\partial^2 p(x_i)}{\partial a_j \partial a_k} \simeq N \int \frac{\partial^2 p(x)}{\partial a_j \partial a_k} dx = N \frac{\partial^2}{\partial a_j \partial a_k} \int p(x) dx = 0$$

da die Normierung $\int p(x; \vec{a}) dx = 1$ unabhängig von den Parametern \vec{a} ist. Der zweite Term der Gl. 5.39 gibt in dieser Näherung eine Abschätzung der Kovarianzmatrix $C = 1/S$

$$S_{jk} = - \frac{\partial^2 \ln L}{\partial a_j \partial a_k} \simeq N \int \frac{1}{p(x)} \frac{\partial p(x)}{\partial a_j} \frac{\partial p(x)}{\partial a_k} \quad (5.40)$$

Man vergleiche diese a priori Abschätzung der Fehler ("magische" Formel) mit der a posteriori Berechnung der Fehler mit der Gl. 5.38. Der S -Matrix ist proportional zur Zahl der zu messenden Ereignisse N und die Fehler der Parameter damit proportional $1/\sqrt{N}$.

Als Beispiel betrachten wir wieder den Gauss-Peak über dem exponentiellen Untergrund (Kap. 5.3). Die Theoriefunktion ist (Gl. 5.28)

$$p(x; r) = r g(x) + (1 - r) e(x)$$

wobei $g(x)$ eine normierte Gauss-Funktion ist ($x_0 = 6$, $\sigma = 0.7$) und $e(x)$ eine normierte Exponentialfunktion ($\lambda = 0.5$). Es gibt also nur einen Parameter $r = N_g/N$, wobei N_g die Zahl der Ereignisse im Gauss-Peak ist und N die Gesamtzahl der Ereignisse. Die Ableitung $dp(x)/dr$ ist trivial und das Integral (Gl. 5.40) kann numerisch berechnet werden (Program 5.7). Das Resultat für $r = 1/3$ ist

$$\sigma_r = 0.534/\sqrt{N} = 0.0195 \quad \text{für } N = 746$$

in Übereinstimmung mit dem a posteriori Fehler 0.0196 (Gl. 5.31). Für $x_0 = 16$ sind Gauss-Peak und exponentieller Untergrund praktisch getrennt und der Fehler ist $0.472/\sqrt{N}$ in Übereinstimmung mit dem Fehler der Binomialverteilung $\sqrt{r(1-r)/N} = \sqrt{2/9N} = 0.471/\sqrt{N}$.

```
%-----
%   MATLAB program 5.7: error estimation with magic formula
%-----
tau = 2;          sig = 0.7;
x0 = 6;          dx = 0.1;
x = dx/2:dx:20-dx/2;
ex = (dx/tau)*exp(-x/tau);          % normalized exponential
gx = dx*gauss(x,x0,sig);          % normalized gaussian
r = 1/3;          px = r*gx + (1-r)*ex;
dpdr2 = (gx - ex).^2;          S = sum(dpdr2 ./ px);
sigma_r = sqrt(1/S);
%-----
```