

# Parameterschätzung für Poisson-Verteilungen bei unvollständiger Beobachtung

Michael Wodny  
Institut für Biometrie und Medizinische  
Informatik  
W.-Rathenau-Str. 48  
17475 Greifswald  
wodny@biometrie.uni-greifswald.de

## Zusammenfassung

Betrachtet werden die Schätzungen des Parameters  $\lambda$  einer Poisson-Verteilung.  $\lambda$  entspricht der erwarteten Anzahl von Objekten in einer definierten Probe. Die unvollständige Beobachtung resultiert daraus, dass in den einzelnen Proben die interessierenden Objekte nicht direkt gezählt werden können. Durch entsprechende Experimente kann nur entschieden werden, ob die Objekte vorhanden waren oder nicht. Solche Problemstellungen findet man z. B. in der Medizin und Biologie. Die Schätzung von  $\lambda$  ergibt sich aus der experimentellen Auswertung definierter Verdünnungsreihen. Es wird nachgewiesen, dass die bekannte ML-Schätzung eindeutig bestimmt ist. Für die Minimum- $\chi^2$ -Schätzung gelingt das nicht. Es ergibt sich aber, dass die variierte Minimum- $\chi^2$ -Methode und die ML-Schätzung übereinstimmen.

Für die Schätzungen können asymptotische Konfidenzintervalle angegeben werden. In der Literatur wird empfohlen, nicht  $\lambda$  direkt, sondern  $\mu := \ln(\lambda)$  zu schätzen, weil deren Verteilung einer Normalverteilung "ähnlicher" sieht. In einem Simulationsexperiment wurde festgestellt, dass beide Schätzungen das  $\alpha$ -Risiko einhalten. Die vorgegebenen  $\lambda$  variierten zwischen 1 und 30.

Als rechentechnische Lösung wird ein SAS-Programm vorgestellt. Durch die Verwendung von Arrays kann die vorgeschlagene Lösung mit dem Standardumfang von SAS genutzt werden.

**Schlüsselwörter:** Poisson-Verteilung, ML-Schätzung, Simulationsexperiment.

## 1 ML-Schätzung und eindeutige Lösbarkeit

In den biologischen und medizinischen Wissenschaften tritt häufig folgendes Schätzproblem auf: In einer größeren Probenmenge soll die Anzahl gewisser Teilchen pro Volumeneinheit bestimmt werden. Das können Mikroorganismen im Boden oder Krebszellen in Körperflüssigkeiten sein. Das Problem besteht darin, dass die zu zählenden Objekte nicht direkt beobachtbar sind. Durch bestimmte Duplizierungsverfahren, Anzüchtungsversuche oder chemische Tests ist nur entscheidbar, ob sich in der Volumeneinheit die gesuchten Objekte befinden oder nicht. Es ist bei einem positiven Ausgang des Experimentes nicht zu sagen, wie viele es gewesen sind. Der Test ist qualitativ, nicht quantitativ.

Ein konkretes Beispiel ist die Knochenmarkstransplantation bei bösartigen Bluterkrank-

kungen. Nachdem den Patienten durch eine Chemotherapie das eigene blutbildende Organ zerstört wurde, wird Knochenmark eines Spenders infundiert. Die Stammzellen siedeln sich an und bauen ein neues System auf. Wenn kein passender Spender gefunden werden kann, sind Eigenspenden eine mögliche Alternative. Das gewonnene Knochenmark wird extern so behandelt, dass möglichst wenig krankhafte Zellen überleben. Um abzuschätzen, wie viele Krebszellen sich noch in dem zur Transplantation vorgesehenen Material befinden, werden definierte Proben entnommen und diese so lange vermehrt, bis eventuelle Krebszellen nachweisbar sind.

Um eine Schätzung der Anzahl zu ermöglichen, wird eine definierte Verdünnungsreihe des Materials erzeugt und mit jeder Verdünnungsstufe eine gewisse Anzahl der Duplikationsversuche oder Tests durchgeführt. Wenn die dafür verwendete Menge in jeder Stufe konstant eine Volumeneinheit bleibt, verringert sich die Wahrscheinlichkeit dafür, dass sich die gesuchten Objekte in den Probenvolumina befinden in dem Verhältnis der Verdünnung und die Anzahl der positiven Testergebnisse in den Stufen geht gegen Null mit wachsender Verdünnung.

Grundvoraussetzung zur Lösung des Ausgangsproblems ist, dass die Anzahl der zu bestimmenden Objekte gegenüber den anderen Teilchen der Probe sehr klein ist. Damit gehorcht deren Anzahl  $X$  pro Volumeneinheit einer Poisson-Verteilung mit dem Parameter  $\lambda$ . Es gilt

$$P(X = j) = \frac{\lambda^j}{j!} e^{-\lambda}, j=0,1,2,3 \dots$$

Damit ist beschrieben, mit welcher Wahrscheinlichkeit sich im fixierten Volumen genau  $j$  der gesuchten Objekte befinden. Da der Erwartungswert einer poissonverteilten Zufallsgröße gleich  $\lambda$  ist, entspricht  $\lambda$  der gesuchten mittleren Anzahl von Objekten pro Volumeneinheit in der Ausgangsprobe.

Wie schon ausgeführt, kann aufgrund der Versuchsbedingungen nur entschieden werden, ob sich in den Probentöpfchen einer Volumeneinheit Objekte befunden haben oder nicht. Die Antwort ist negativ mit der Wahrscheinlichkeit  $e^{-\lambda}$  (das entspricht  $P(X=0)$ ) und entsprechend positiv mit  $1-e^{-\lambda}$  ( $P(X \neq 0)$ ). Es soll angenommen werden, dass  $a$  Proben genommen worden sind und unter diesen  $k$  positiv reagiert haben. Bei einem Binomialansatz ist die Wahrscheinlichkeit dafür, dass gerade  $k$  Versuche positiv ausgehen

$$P(B = k) = \binom{a}{k} (1 - e^{-\lambda})^k (e^{-\lambda})^{(a-k)}$$

Wenn angenommen wird, dass  $n-1$  Verdünnungsstufen  $V_i$ ,  $i=2, \dots, n$ , angefertigt werden, die zur Ausgangskonzentration  $V_1$  im Verhältnis  $V_1/V_i =: v_i$  stehen, in jeder Stufe  $a_i$  Versuche durchgeführt werden und davon  $k_i$  positiv ausgehen, gilt für jede Stufe

$$P(B_i = k_i) = \binom{a_i}{k_i} (1 - e^{-\lambda/v_i})^{k_i} (e^{-\lambda/v_i})^{(a_i - k_i)}, i=1,2, \dots, n.$$

In der Praxis werden im Allgemeinen die Verdünnungsstufen in Zweierpotenzen ange-

legt, d. h. von einer Stufe zur nächsten wird auf die Hälfte verdünnt, sodass sich in den Volumeneinheiten im Mittel halb so viele Objekte befinden wie in der vorhergehenden Stufe. Um die Argumentation nicht unnötig kompliziert zu machen, soll von so einer Verdünnungsreihe ausgegangen werden. Es gilt dann  $v_i=2^{i-1}$ ,  $i=1,2,\dots,n$ . Die Argumentation gilt auch für beliebige  $v_i$ . Wichtig ist die Feststellung, dass zur Erzeugung jeder Verdünnungsstufe immer wieder auf das Ausgangsmaterial zurückgegriffen wird.

Um zu erkennen, dass die Versuche in der gesamten Verdünnungsreihe stochastisch unabhängig sind, denke man sich ein kleinstes Schöpfgefäß von  $1/2^{n-1}$  Volumeneinheiten. Mit diesem wird immer wieder *aus dem Ausgangsmaterial* entnommen. Da eine  $2^{n-1}$ -fache Verdünnung auch mit diesem kleinsten Topf realisiert werden kann, werden  $a_n$  Probenentnahmen benötigt um die Versuche der letzten Verdünnungsstufe zu realisieren. In der davor liegenden Stufe ist die Konzentration im Mittel doppelt so hoch, sodass für jede der  $a_{n-1}$  Proben je zwei Inhalte des kleinsten Schöpfgefäßes benötigt werden. Eine Probe in der  $(n-1)$ -ten Stufe zeigt nur dann ein negatives Resultat, wenn sich in beiden Inhalten kein Objekt befindet:

$$P(B_n = 0) P(B_n = 0) = e^{-\lambda/2^{n-1}} e^{-\lambda/2^{n-1}} = e^{-\lambda/2^{n-2}} = P(B_{n-1} = 0).$$

Damit sind die einzelnen Probenentnahmen unabhängig.

Der Parameter  $\lambda$  wird so bestimmt, dass die Anzahlen der konkret aufgetretenen positiven Ereignisse  $k_1, k_2, \dots, k_n$  maximale Wahrscheinlichkeit besitzt. Aufgrund der stochastischen Unabhängigkeit bedeutet das, dass die Likelihoodfunktion

$$L(\lambda) := \prod_{i=1}^n \binom{a_i}{k_i} (1 - e^{-\lambda/2^{i-1}})^{k_i} (e^{-\lambda/2^{i-1}})^{(a_i - k_i)}$$

maximal gemacht werden muss. Da der natürliche Logarithmus  $\ln$  eine streng monotone Funktion ist, kann anstelle von  $L(\lambda)$  auch die einfacher zu behandelnde Funktion

$$\ln(L(\lambda)) = \sum_{i=1}^n \left[ \ln \binom{a_i}{k_i} + k_i \ln(1 - e^{-\lambda/2^{i-1}}) - (a_i - k_i) \frac{\lambda}{2^{i-1}} \right] \tag{1}$$

maximiert werden. Diese ist nur für  $\lambda > 0$  definiert.

Notwendige Bedingung für ein lokales Extremum ist

$$\frac{d \ln(L(\lambda))}{d \lambda} = \ln(L(\lambda))' = \sum_{i=1}^n \left[ \frac{k_i}{2^{i-1} (e^{\lambda/2^{i-1}} - 1)} - \frac{a_i - k_i}{2^{i-1}} \right] = 0.$$

Diese Funktion ist nur für  $\lambda=0$  nicht definiert. Für positive  $\lambda$  gilt

$$\lim_{\lambda \rightarrow 0^+} \ln(L(\lambda))' = +\infty \quad \text{bzw.} \quad \lim_{\lambda \rightarrow +\infty} \ln(L(\lambda))' = \sum_{i=1}^n \left[ -\frac{a_i - k_i}{2^{i-1}} \right] \leq 0$$

und für (nicht interessierende) negative

$$\lim_{\lambda \rightarrow 0^-} \ln(L(\lambda))' = -\infty, \quad \lim_{\lambda \rightarrow -\infty} \ln(L(\lambda))' = \sum_{i=1}^n \left[ -\frac{a_i}{2^{i-1}} \right] < 0.$$

Weil die zweite Ableitung

$$\frac{d^2 \text{Ln}(L(\lambda))}{d\lambda^2} = \text{Ln}(L(\lambda))'' = -\sum_{i=1}^n \left[ \frac{k_i e^{\lambda/2^{i-1}}}{(2^{i-1})^2 (e^{\lambda/2^{i-1}} - 1)^2} \right]$$

kleiner als Null ist für alle  $\lambda \neq 0$ , ist  $\text{Ln}(L(\lambda))'$  streng monoton fallend für *alle*  $\lambda \neq 0$ . Damit existiert *genau eine* Nullstelle von  $\text{Ln}(L(\lambda))'$  und ein einziges lokales Maximum  $\lambda^*$  der Maximum-Likelihood-Funktion  $L(\lambda)$  bzw.  $\text{Ln}(L(\lambda))$ . Dieses ist größer als Null, wenn wenigstens ein  $k_i < a_i$  existiert.

$\lambda^* = \lambda_{ML}$  kann z. B. mit dem Newton-Verfahren iterativ bestimmt werden:

$$\lambda_{i+1} := \lambda_i - \frac{\text{Ln}(L(\lambda_i))'}{\text{Ln}(L(\lambda_i))''}, \quad i=0,1,2,\dots,m.$$

Als Startwert  $\lambda_0$  sollten kleine Werte benutzt werden. Bewährt hat sich  $\lambda_0=0.01$ . Als zusätzliche Sicherheit können während des Iterationsprozesses auftretende negative  $\lambda_{i+1}$  durch den Betrag ersetzt werden. Große Startwerte führen aufgrund der Gestalt von  $\text{Ln}(L(\lambda))'$  leichter zur Divergenz (vergleiche Abbildung 2).

Die asymptotische Varianz  $V(\lambda_{ML})$  des Maximum-Likelihood-Schätzers ergibt sich zu

$$V(\lambda_{ML}) = -\frac{1}{\text{Ln}(L(\lambda_{ML}))''} \quad (\text{FISHER 1922}).$$

Damit kann ein approximatives  $(1-\alpha)$ -Konfidenzintervall durch die Beziehung

$$\lambda_{ML} \pm z_{1-\alpha/2} \sqrt{V(\lambda_{ML})}$$

angegeben werden.

Eine andere Möglichkeit ist die Beschreibung des Versuchsaufbaus durch ein Polynomialmodell. Das Versuchsergebnis entspricht einer Stichprobe, bei der  $k_i$  Ereignisse mit den Wahrscheinlichkeiten  $(1 - e^{-\lambda/2^{i-1}})$  und  $(a_i - k_i)$  Ereignisse mit den

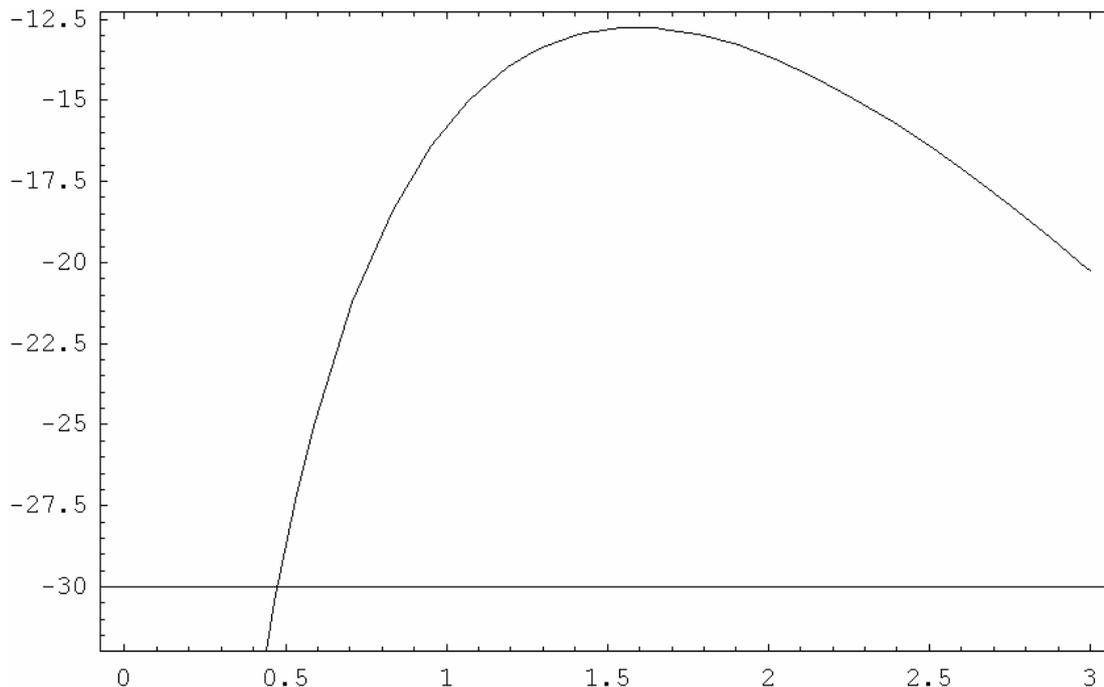
Wahrscheinlichkeiten  $e^{-\lambda/2^{i-1}}$ ,  $i=1,2,\dots,n$ , auftreten. Der Stichprobenumfang entspricht

$N = \sum_{i=1}^n a_i$ . Da der Polynomialkoeffizient unabhängig von  $\lambda$  ist, erhält man den gleichen

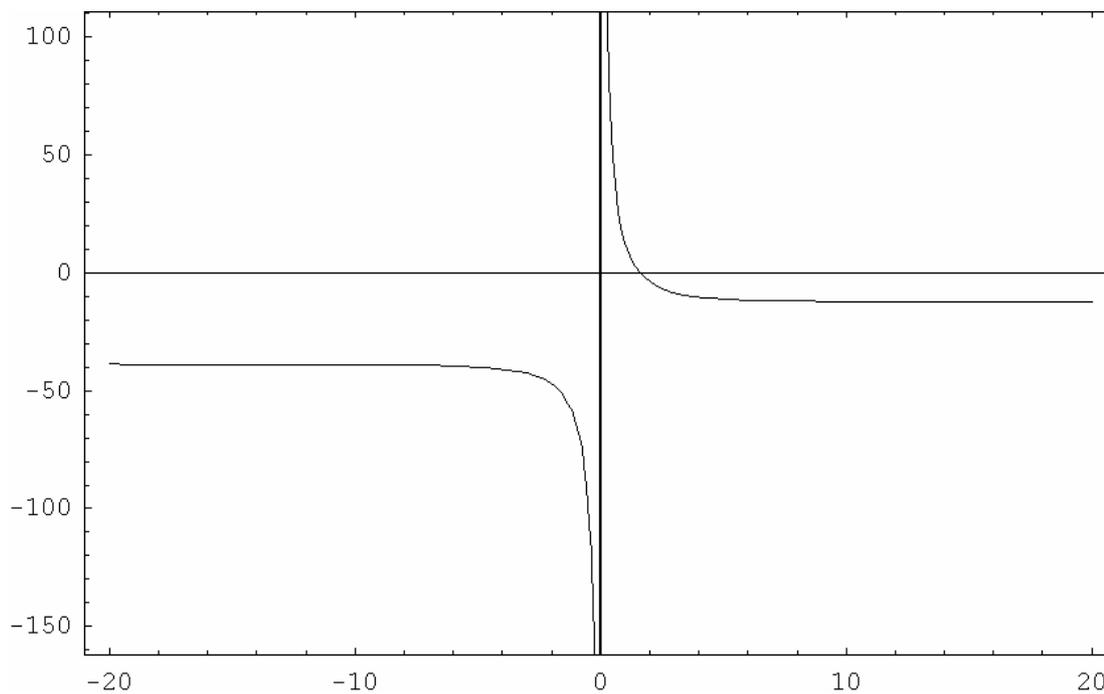
ML-Schätzer wie beim obigen Binomialmodell.

## 2 Beispiel

Als Beispiel sei eine Verdünnungsreihe mit  $n=5$  Stufen,  $v_i=2^{i-1}$  und  $a_i=20$  für alle  $i$  von 1 bis 5 verwendet. Die Anzahlen der positiven Versuchsausgänge seien  $k_1=20$ ,  $k_2=10$ ,  $k_3=5$ ,  $k_4=1$  und  $k_5=0$ . Die erste Abbildung zeigt einen Ausschnitt der Funktion  $\text{Ln}(L(\lambda))$ .



**Abbildung 1:** Ausschnitt der Funktion  $\text{Ln}(L(\lambda))$  zu den Beispieldaten



**Abbildung 2:** Ausschnitt der Ableitung  $\text{Ln}(L(\lambda))'$

Die Schätzung liefert  $\lambda_{\text{ML}} = 1.5896$ . Als Varianz erhält man  $V(\lambda_{\text{ML}}) = 0.2824$  und somit ergibt sich ein approximatives 0.95-Konfidenzintervall von  $[1.0362, 2.1430]$ .

### 3 Simulation

Ziel des Simulationsexperimentes ist der Vergleich des ML-Schätzers  $\lambda_{ML}$  mit einer Modifikation. Es wird nicht  $\lambda$  direkt, sondern  $\mu := \text{Ln}(\lambda)$  geschätzt. Dieses Vorgehen wird vorgeschlagen, weil die Verteilung der  $\mu_{ML}$  einer Normalverteilung "ähnlicher" sieht als die der  $\lambda_{ML}$  (vergleiche Facekas de St. Groth 1982 und Macken 1999).

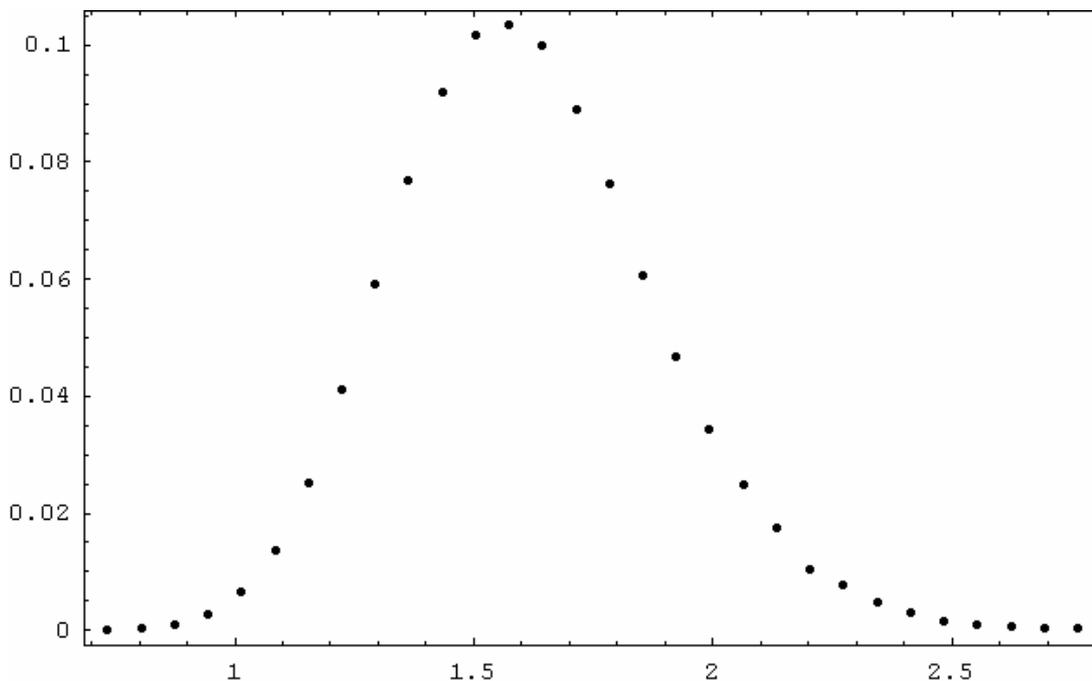
Die Beziehung (1) geht über in

$$\text{Ln}(\mathcal{L}(\mu)) = \sum_{i=1}^n \left[ \text{Ln} \left( \frac{a_i}{k_i} \right) + k_i \text{Ln}(1 - \exp(-e^\mu / 2^{i-1})) - (a_i - k_i) \frac{e^\mu}{2^{i-1}} \right].$$

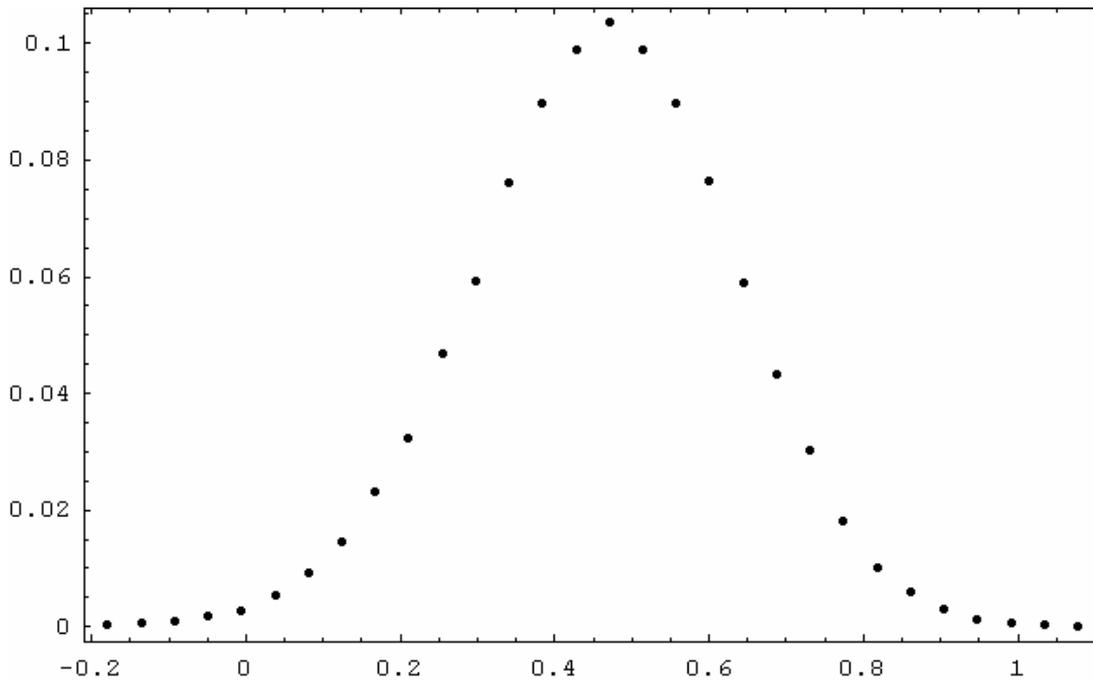
Es sei  $\mu_{ML}$  der ebenfalls eindeutig bestimmte Maximum-Likelihood-Schätzer für  $\mu$ . Dieser besitzt die asymptotische Varianz  $\text{Var}(\mu_{ML}) = -1/\text{Ln}(\mathcal{L}(\mu_{ML}))''$  und das approximative  $(1-\alpha)$ -Konfidenzintervall  $G_{\pm} := \mu_{ML} \pm z_{1-\alpha/2} \cdot \sqrt{\text{Var}(\mu_{ML})}$

Der eigentlich gesuchte Schätzer  $\lambda$  folgt aus  $\lambda = \lambda_{ML} = e^{\mu_{ML}}$  und dessen approximatives Konfidenzintervall aus

$$[ e^{G^-} , e^{G^+} ]. \tag{2}$$



**Abbildung 3:** Verteilung der geschätzten Anzahlen  $\lambda_{ML}$  bei einem Simulationsexperiment mit  $\lambda = 1.59$ ,  $n = 5$ ,  $a_i = 20$ ,  $i=1,2,\dots,5$ , und einem Simulationsumfang von  $10^5$ .



**Abbildung 4:** Verteilung der geschätzten  $\mu_{ML}$  mit  $\mu = \text{Ln}(\lambda) = \text{Ln}(1.59) = 0.464$  bei einem analogen Simulationsexperiment.

In Tabelle 1 sind bei variierten vorgegebenen  $\lambda$  und ansonsten unveränderten Simulationsbedingungen mit  $n=5$  und  $a_i=20, i=1,2,\dots,n$ , und  $10^5$  Runs unter anderem die ermittelten  $\alpha$ -Risiken aufgeführt, mit dem das vorgegebene  $\lambda$  nicht im berechneten approximativen 0.95-Konfidenzintervall enthalten war. Man beachte, dass bei Beibehaltung von  $n$  und  $a_i$  mit wachsendem  $\lambda$  die Wahrscheinlichkeit für  $k_i=a_i$  besonders bei kleinen  $i$  wächst. Konsequenz ist eine größere empirische Standardabweichung. Deshalb ist bei  $\lambda=20$  und  $\lambda=30$   $n=10$  verwendet worden.

**Tabelle 1:** Simulationsergebnisse bei der ML-Schätzung von  $\lambda$ .

Vorgegebenes $\lambda$	Mittelwert der $\lambda_{ML}$ bei der Simulation	empirische Standardabweichung der $\lambda_{ML}$	$\alpha$ -Risiko
1.59	1.611	0.2732	0.0528
3.00	3.051	0.4765	0.0497
5.00	5.102	0.8021	0.0483
10.00	10.262	1.7460	0.0488
20.00	20.330	2.9834	0.0500
30.00	30.510	4.5116	0.0500

Das  $\alpha$ -Risiko wird in den hier untersuchten Bereichen von  $\lambda$  auch bei der Schätzung von  $\lambda_{ML}$  im Wesentlichen eingehalten.

In Tabelle 2 sind bei den gleichen vorgegebenen  $\lambda$  und ansonsten unveränderten Simulationsbedingungen die Ergebnisse der ML-Schätzung von  $\mu=\text{Ln}(\lambda)$  mitgeteilt. Das  $\alpha$ -Risiko entspricht wieder der relativen Häufigkeit, mit der das vorgegebene  $\lambda$  nicht in den über (2) berechneten 0.95-Konfidenzintervallen enthalten war.

**Tabelle 2:** Simulationsergebnisse bei der ML-Schätzung von  $\mu$

Vorgegebenes $\lambda$	Mittelwert $\mu_{\text{mit}}$ der $\mu_{\text{ML}}$	$\lambda_{\text{mit}}=e^{\mu_{\text{mit}}}$	empirische Standardabweichung der $\mu_{\text{ML}}$	$\alpha$ -Risiko
1.59	0.4628	1.5885	0.1703	0.0491
3.00	1.1030	3.0133	0.1557	0.0492
5.00	1.6180	5.0415	0.1556	0.0415
10.00	2.3152	10.1274	0.1673	0.0500
20.00	3.0007	20.1000	0.1464	0.0500
30.00	3.4051	30.1173	0.1468	0.0485

Das  $\alpha$ -Risiko wird eingehalten. Es scheint, dass  $\lambda=e^{\mu}$  ein besserer Schätzer ist, weil  $\lambda_{\text{mit}}=e^{\mu_{\text{mit}}}$  näher am vorgegebenen  $\lambda$  liegt. Das wird aber durch das Logarithmieren vorgetäuscht. Berechnet man bei jedem Simulationsschritt aus  $\mu_{\text{ML}}$  das geschätzte  $\lambda$ , ergibt die Auswertung analoge Ergebnisse wie in Tabelle 1, denn zu vorgegebenen Ausgangsdaten ist  $\lambda_{\text{ML}}=e^{\mu_{\text{ML}}}$ . Die Punktschätzer stimmen überein.

## 4 Andere Schätzmethoden

### 4.1 Die Minimum- $\chi^2$ -Methode

Ausgangspunkt dieser Schätzung von  $\lambda$  ist das PEARSON'sche

$$\begin{aligned} \chi^2 = \chi^2(\lambda) &= \sum_{i=1}^n \left[ \frac{\left(k_i - a_i(1 - e^{-\lambda/2^{i-1}})\right)^2}{a_i(1 - e^{-\lambda/2^{i-1}})} + \frac{\left(a_i - k_i - a_i e^{-\lambda/2^{i-1}}\right)^2}{a_i e^{-\lambda/2^{i-1}}} \right] \\ &= \sum_{i=1}^n \left[ \frac{\left(a_i - k_i - a_i e^{-\lambda/2^{i-1}}\right)^2}{a_i(1 - e^{-\lambda/2^{i-1}}) e^{-\lambda/2^{i-1}}} \right] \end{aligned}$$

Notwendige Bedingung für ein lokales Extremum ist

$$\chi^2(\lambda)' = \sum_{i=1}^n \frac{e^{\lambda/2^{i-1}} (-a_i + (a_i - k_i)e^{\lambda/2^{i-1}})(2k_i - a_i + (a_i - k_i)e^{\lambda/2^{i-1}})}{2^{i-1} a_i (e^{\lambda/2^{i-1}} - 1)^2} = 0.$$

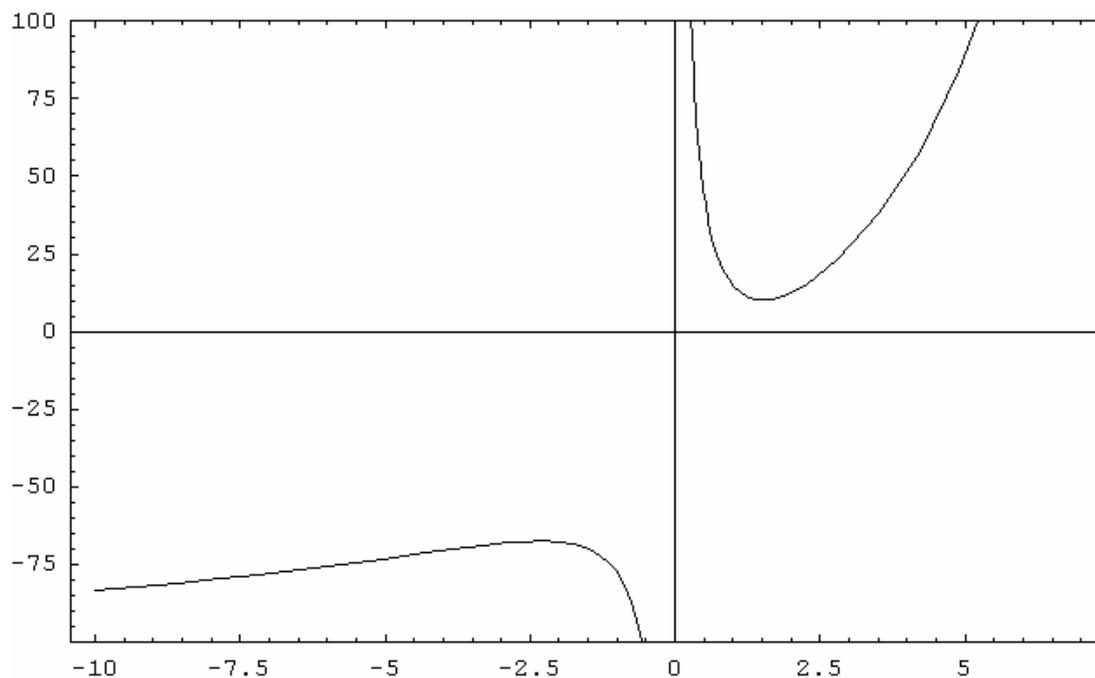


Abbildung 5:  $\chi^2(\lambda)$  für die Beispieldaten

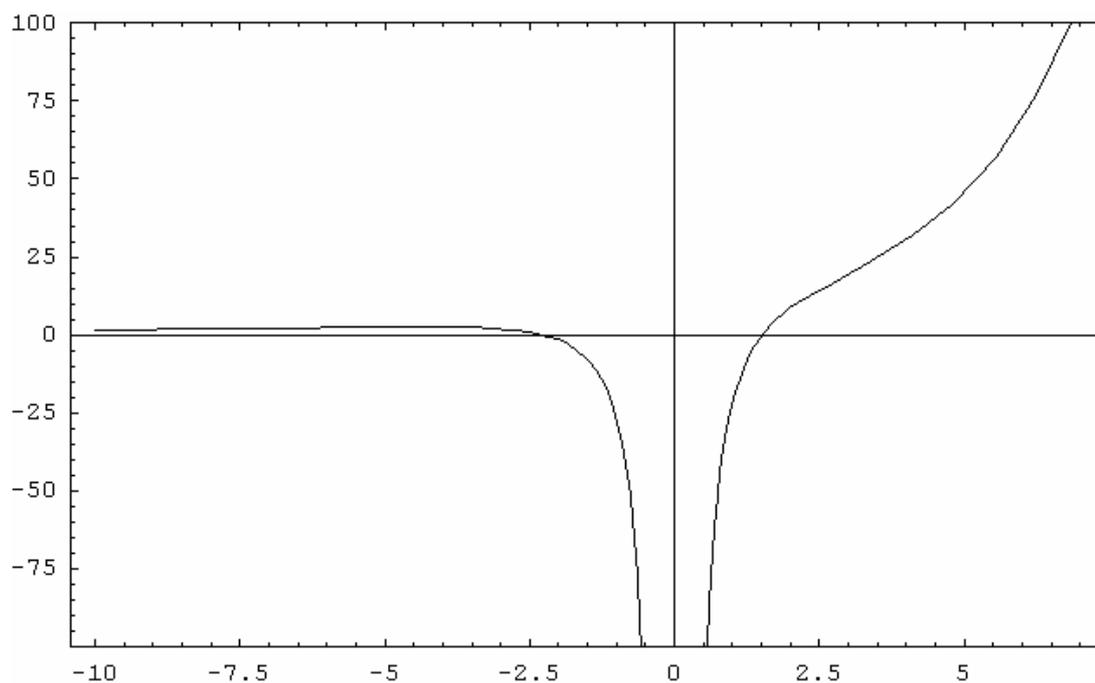


Abbildung 6: Die Ableitung  $\chi^2(\lambda)'$ .

Die eindeutige Lösung der Gleichung  $\chi^2(\lambda)' = 0$  für  $\lambda > 0$  lässt sich allein aus den Monotonieeigenschaften der zweiten Ableitung nicht für den allgemeinen Fall herleiten.

## 4.2 Die variierte Minimum- $\chi^2$ -Methode

Bei der variierten Minimum- $\chi^2$ -Methode wird nur der Zähler der Summanden von  $\chi^2 = \chi^2(\lambda)$  abgeleitet und der Nenner als konstant betrachtet. Es entsteht die Funktion

$$\text{Vari}\chi(\lambda) = \sum_{i=1}^n 2 \left( \frac{a_i(e^{\lambda/2^{i-1}} - 1) - k_i e^{\lambda/2^{i-1}}}{2^{i-1}(e^{\lambda/2^{i-1}} - 1)} \right),$$

deren Nullstellen berechnet werden müssen. Nach gleichzeitiger Addition und Subtraktion von  $k_i$  im Zähler erhält man

$$\begin{aligned} \text{Vari}\chi(\lambda) &= 2 \sum_{i=1}^n \left( \frac{a_i(e^{\lambda/2^{i-1}} - 1) - k_i + k_i - k_i e^{\lambda/2^{i-1}}}{2^{i-1}(e^{\lambda/2^{i-1}} - 1)} \right) \\ &= 2 \sum_{i=1}^n \left( \frac{-k_i}{2^{i-1}(e^{\lambda/2^{i-1}} - 1)} + \frac{a_i - k_i}{2^{i-1}} \right). \end{aligned}$$

Die Nullstellen dieser Funktion stimmen mit denen von  $\text{Ln}(L(\lambda))'$  überein. Daraus folgt, dass bei der vorliegenden Problemstellung die Maximum-Likelihood-Schätzung und die Schätzung nach der variierten Minimum- $\chi^2$ -Methode *identisch* sind (vergleiche auch Rasch, Seite 223).

## 5 SAS-Lösung

```
%LET Anz_Verd=5;      * Anzahl der Verdünnungen oder Dimension der
                        Vektoren a, k und VR;
DATA ERG(KEEP = Lambda Abl1 untere_Konf_grenze obere_Konf_grenze);
  ARRAY a{*} a1-a&Anz_Verd; ARRAY k{*} k1-k&Anz_Verd;
  ARRAY VR{*} VR1-VR&Anz_Verd;
  INPUT a{*} k{*} VR{*};
  n=DIM(a); Lambda=0.01; D1=0;
  WEI:Abl1=0; Abl2=0;
  DO i=1 TO n; e=Exp(Lambda/VR[i]);
    Abl1=Abl1 + k[i]/VR[i]/(e-1) - (a[i]-k[i])/VR[i];
    Abl2=Abl2 - k[i]*e/(VR[i]*(e-1))**2;
  END;
  D1=D1+1; IF D1>50 THEN GOTO AU;
  Lamneu=ABS(Lambda - Abl1/Abl2);
  IF ABS(Lambda - Lamneu)<1E-12 THEN GOTO AU;
  Lambda=Lamneu; GOTO WEI;
  AU:StAbw=SQRT(-1/Abl2);
  untere_Konf_grenze=Lambda - 1.95996*StAbw;
```

```

    obere_Konf_grenze =Lambda + 1.95996*StAbw;
OUTPUT;
/*****/
/* Daten der Verdünnungsversuche */
/* 1. Zeile: Anzahl der Versuche a[i] in den Verdünnungsreihen */
/* 2. Zeile: Anzahl der positiven Ergebnisse k[i] von a[i] */
/* 3. Zeile: Verdünnungsfaktoren VR[i] */
/*****/
DATALINES;
  20  20  20  20  20
  20  10  5  1  0
  1  2  4  8  16
;
RUN;
TITLE1 "ML-Schätzung der mittleren Anzahl Lambda aus
Verdünnungsstufen";
TITLE2 "mit zugehörigem approximativem 0.95-Konfidenzintervall.";
PROC PRINT DATA=ERG NOOBS; RUN;
PROC DELETE DATA=ERG; RUN;
TITLE1; TITLE2;

```

### 5.1 Programmausgabe

ML-Schätzung der mittleren Anzahl Lambda aus Verdünnungsreihen mit zugehörigem approximativem 0.95-Konfidenzintervall. 1				
			untere_	obere_
			Konf_	Konf_
			grenze	grenze
Lambda	Ab11			
1.58959	2.7467E-12		1.03617	2.14300

Lambda: ML-Schätzung für  $\lambda$

Ab11:  $\ln(L(\text{Lambda}))'$   
 Liegt Ab11 nicht nahe bei Null, dann kann es sich **nicht** um die Lösung handeln. Das Newton-Verfahren hat bei den vorgesehenen 50 Iterationen keine konvergente Folge erzeugt.

## **Literatur**

- [1] de St. Fazekas, S. Groth,: The evaluation of limiting dilution assays. *J. Immunol. Methods*, 49, 11, 1982.
- [2] R.A. Fisher: On the mathematical foundations of theoretical statistics. *Philos. Trans. R. Soc. London, A.* 222-309, 1922.
- [3] C. Macken: Design and analysis of serial limiting dilution assays with small sample sizes. *Journal of Immunological Methods* 222, 13-29, 1999.
- [4] D. Rasch: Einführung in die mathematische Statistik Band I  
Deutscher Verlag der Wissenschaften Berlin, 1978.
- [5] C. Taswell: Limiting dilution assays for the separation, characterization, and quantitation of biologically active particles and their clonal progeny  
In: T.G. Pretlow, T.P. Pretlow (Eds):  
*Cell Separation: Methods and Selected Applications*, Vol 4, Academic Press, New York, 1987
- [6] SAS Institute Inc.: SAS Language: Reference Version 6, First Edition. Cary, NC 27513, 1990.