Technische Universität München
Fakultät für Informatik
Lehrstuhl für Theoretische Informatik
Prof. Dr. Susanne Albers
Dennis Kraft, Richard Stotz

Sommersemester 2015
Lösungsblatt 10
6. Juli 2015

Diskrete Wahrscheinlichkeitstheorie

Tutoraufgabe 1

Wir betrachten eine beliebige Konvexkombination $Y = \sum_{i=1}^{n} \lambda_i X_i$ über n unabhängigen Stichprobenvariablen von X.

1. Zeigen Sie, dass es sich bei Y um einen erwartungstreuen Schätzer für $\mathbb{E}[X]$ handelt.
2. Bestimmen Sie Koeffizienten λ_i, welche die Effizienz von Y optimieren. Beweisen Sie Ihre Konstruktion.

Hinweis: Zeigen Sie in der zweiten Teilaufgabe zunächst, dass die gewichtete Summe $\frac{1}{n} \cdot \sum_{i=1}^{n} \lambda_i^2$ größer gleich $\frac{1}{n^2}$ ist.

Lösungsvorschlag

Nachdem es sich bei Y um eine Konvexkombination handelt, summieren sich die Koeffizienten λ_i zu 1 auf. Aus der Linearität des Erwartungswertes folgt somit unmittelbar, dass es sich bei Y um einen erwartungstreuen Schätzer für $\mathbb{E}[X]$ handelt

$$\mathbb{E}[Y] = \mathbb{E}\left[\sum_{i=1}^{n} \lambda_i X_i\right] = \sum_{i=1}^{n} \lambda_i \mathbb{E}[X_i] = \sum_{i=1}^{n} \lambda_i \mathbb{E}[X] = \mathbb{E}[X] \sum_{i=1}^{n} \lambda_i = \mathbb{E}[X].$$

Betrachten wir als Nächstes die Effizienz von Y. Um diese zu optimieren, müssen wir die Koeffizienten λ_i so wählen, dass die mittlere quadratische Abweichung von Y, in unserem Fall also die Varianz von Y da

$$\mathbb{E}[(Y - \mathbb{E}[X])^2] = \mathbb{E}[(Y - \mathbb{E}[Y])^2] = \mathrm{Var}[Y]$$

gilt, minimal wird. Hierzu definieren wir eine weitere Zufallsvariable Z, welche jeden Wert λ_i jeweils mit Wahrscheinlichkeit $\frac{1}{n}$ annimmt. Nachdem Z lediglich abzählbar viele Werte annimmt, wissen wir, dass sowohl der Erwartungswert als auch die Varianz von Z existieren müssen. Insbesondere gilt somit $\mathbb{E}[Z^2] \geq \mathbb{E}[Z]^2$, was wir wiederum nutzen können um den Hinweis aus der Aufgabenstellung zu beweisen

$$\frac{1}{n} \cdot \sum_{i=1}^{n} \lambda_i^2 = \mathbb{E}[Z^2] \geq \mathbb{E}[Z]^2 = \left(\frac{1}{n} \cdot \sum_{i=1}^{n} \lambda_i\right)^2 = \frac{1}{n^2}.$$

Kommen wir nun zur Varianz von Y. Nachdem die Stichporbenvariablen X_i unabhängig sind gilt

$$\mathrm{Var}[Y] = \mathrm{Var}\left[\sum_{i=1}^{n} \lambda_i X_i\right] = \sum_{i=1}^{n} \lambda_i^2 \mathrm{Var}[X_i] = \sum_{i=1}^{n} \lambda_i^2 \mathrm{Var}[X] = \mathrm{Var}[X] \cdot \sum_{i=1}^{n} \lambda_i^2.$$

Mit dem soeben bewiesenen Hinweis aus der Angabe, können wir diesen Term nach unten abschätzen zu

$$\mathrm{Var}[X] \cdot \sum_{i=1}^{n} \lambda_i^2 \geq \frac{\mathrm{Var}[X]}{n}.$$

Wählen wir für jedes λ_i den Wert $\frac{1}{n}$, so gilt

$$\mathrm{Var}[Y] = \mathrm{Var}[X] \cdot \sum_{i=1}^{n} \frac{1}{n^2} = \frac{\mathrm{Var}[X]}{n}$$

und die Ungleichung ist mit Gleichheit erfüllt. Folglich optimieren wir die Effizienz von Y, wenn wir jede Stichprobenvariable gleich gewichten, was auch unserer intuitiven Vorstellung entspricht.

Tutoraufgabe 2

Im Rahmen einer wissenschaftlichen Studie fangen Sie n zufällig und unabhängig gewählte Wolpertinger und vermessen diese. Die Größe des i-ten Wolpertingers sei gegeben durch die Stichprobenvariable X_i. Aus früheren Studien wissen Sie, dass die Variablen X_i auf dem Intervall $[0, m]$ gleichverteilt sind, wobei m die noch unbekannte Größe des größten Wolperdingers ist. Ausgehend von ihren Stichproben möchten Sie m abschätzen.

1. Definieren Sie einen Koeffizienten c, so dass das gewichtete Stichprobenmittel $c \sum_{i=1}^{n} \frac{1}{n} X_i$ ein erwartungstreuer Schätzer für m ist.
2. Konstruieren Sie zum Vergleich einen Maximum-Likelihood-Schätzer für m und zeigen Sie, dass dieser nicht erwartungstreu ist. Bestimmen Sie außerdem einen Koeffizienten um dies zu korrigieren.

Lösungsvorschlag

Sei $\bar{X}$ das Stichprobenmittel über X_1 bis X_n. Zur Lösung der ersten Teilaufgabe berechnen wir den Erwartungswert $\mathbb{E}[\bar{X}]$ in Abhängigkeit von m

$$\mathbb{E}[\bar{X}] = \mathbb{E}\left[\sum_{i=1}^{n} \frac{1}{n} X_i\right] = \frac{1}{n} \cdot \sum_{i=1}^{n} \mathbb{E}[X_i] = \frac{1}{n} \cdot \sum_{i=1}^{n} \mathbb{E}[X] = \frac{1}{n} \cdot \sum_{i=1}^{n} \frac{m}{2} = \frac{m}{2}.$$

Skalieren wir das Stichprobenmittel also mit dem Faktor 2, so gilt

$$\mathbb{E}[2\bar{X}] = 2 \cdot \mathbb{E}[\bar{X}] = m,$$

was einem erwartungstreuen Schätzer für m entspricht. Betrachten wir nun den Maximum-Likelihood-Schätzer. Seien hierfür x_1 bis x_n die konkreten Stichprobenwerte, die Sie beobachtet haben, und x das Maximum über die einzelnen x_i. Laut Definition muss der Maximum-Likelihood-Schätzer die Likelihood-Funktion $L(x_1, \ldots, x_n; m)$ bezüglich des Parameters m optimieren. Die Likelihood-Funktion ist wiederum gegeben durch das Produkt der Dichten von x_i. Für alle $m \geq x$ ist die Likelihood-Funktion somit definiert als

$$L(x_1, \ldots, x_n; m) = \prod_{i=1}^{n} f_{X_i}(x_i) = \prod_{1=i}^{n} \frac{1}{m} = \frac{1}{m^n}$$

und für alle anderen Werte von m ist $L(x_1, \ldots, x_n; m)$ automatisch 0. Da die Funktion $\frac{1}{m^n}$ mit steigendem m streng monoton fällt, ist $L(x_1, \ldots, x_n; m)$ am Punkt x maximal. Somit ist unser Maximum-Likelihood-Schätzer, welchen wir im Folgenden mit Y bezeichnen, gegeben durch

$$Y = \max\{X_1, \ldots, X_n\}.$$

Wir berechnen nun den Erwartungswert dieses Schätzers für identische und unabhängige Stichprobenvariablen X_i. Die Verteilungsfunktion von Y ist gegeben durch

$$F_Y(y) = \Pr[X_1 \leq y, \ldots, X_n \leq y] = \prod_{i=1}^{n} \Pr[X_i \leq y] = \begin{cases} 0 & \text{falls } y < 0 \\ \left(\frac{y}{m}\right)^n & \text{falls } 0 \leq y \leq m \\ 1 & \text{falls } m < y \end{cases}.$$

Da diese Funktion überall stetig und differenzierbar ist, bestimmen wir die Dichte durch Ableiten

$$f_Y(y) = F'_Y(y) = \begin{cases} \frac{ny^{n-1}}{m^n} & \text{falls } 0 \leq y \leq m \\ 0 & \text{sonst} \end{cases}$$

und erhalten den Erwartungswert

$$\mathbb{E}[Y] = \int_0^m y \cdot f_Y(y)\mathrm{d}y = \frac{n}{m^n} \cdot \int_0^m y^n \mathrm{d}y = \frac{n}{m^n} \cdot \frac{m^{n+1}}{n+1} = \frac{mn}{n+1}.$$

Folglich ist Y nicht erwartungstreu bezüglich m. Wählen wir als Koeffizienten allerdings $\frac{n+1}{n}$, so ist der Schätzer $\frac{n+1}{n} \cdot Y$ tatsächlich erwartungstreu.

Tutoraufgabe 3

Sei X eine normalverteilte Zufallsvariable mit unbekanntem Erwartungswert μ, einer Standardabweichung von 2 und dem Stichprobenmittel $\bar{X}$ über n unabhängige Stichproben.

1. Leiten Sie eine möglichst kleine untere Schranke für n her, so dass die Wahrscheinlichkeit $\Pr[|\bar{X} - \mu| \leq \frac{1}{10}]$ größer als $\frac{8}{10}$ ist.

2. Sei die Anzahl der Stichproben nun gleich 1600. Geben Sie ein möglichst kleines Konfidenzintervall für μ mit Konfidenzniveau $\frac{8}{10}$ an.

Hinweis: Nutzen Sie die Tabelle der Standardnormalverteilung auf der letzten Seite des Übungsblattes.

Lösungsvorschlag

Aus der Additivität der Normalverteilung können wir schließen, dass das Stichprobenmittel $\bar{X}$ normalverteilt ist mit Erwartungswert μ und Standardabweichung $\frac{2}{\sqrt{n}}$. Transformieren wir $\bar{X}$ nun geeignet, so erhalten wir die standardnormalverteilte Zufallsvariable

$$Y = \sqrt{n}\frac{\bar{X} - \mu}{2}.$$

Als nächstes formen wir die Wahrscheinlichkeit $\Pr[|\bar{X}-\mu| \leq \frac{1}{10}]$ so um, dass sie nur noch von Y und n abhängt

$$\begin{aligned}\Pr\left[|\bar{X}-\mu| \leq \frac{1}{10}\right] &= \Pr\left[-\frac{1}{10} \leq \bar{X}-\mu \leq \frac{1}{10}\right] \\ &= \Pr\left[-\frac{\sqrt{n}}{20} \leq \sqrt{n}\frac{\bar{X}-\mu}{2} \leq \frac{\sqrt{n}}{20}\right] \\ &= \Pr\left[-\frac{\sqrt{n}}{20} \leq Y \leq \frac{\sqrt{n}}{20}\right].\end{aligned}$$

Mit Hilfe der Verteilungsfunktion Φ der Standardnormalverteilung, folgt

$$\Phi\left(\frac{\sqrt{n}}{20}\right) - \Phi\left(-\frac{\sqrt{n}}{20}\right) = 2\Phi\left(\frac{\sqrt{n}}{20}\right) - 1.$$

Da diese Wahrscheinlichkeit größer als $\frac{8}{10}$ sein soll, muss $\Phi(\frac{\sqrt{n}}{20})$ also größer als $\frac{9}{10}$ sein. Wir nutzen nun die Tabelle auf der letzten Seite des Übungsblattes und sehen, dass das $\frac{9}{10}$-Quantil der Standardnormalverteilung mit dem Wert 1,28 angegeben ist. Also approximieren wir $\Phi(1{,}28)$ mit $\frac{9}{10}$ und erhalten eine untere Schranke von

$$n \geq \lceil (20 \cdot 1{,}28)^2 \rceil = 656.$$

Um umgekehrt ein Konfidenzintervall für μ zum Konfidenzniveau $\frac{8}{10}$ anzugeben, müssen wir einen Parameter c bestimmen, so dass $\Pr\left[|\bar{X}-\mu| \leq c\right]$ mindestens $\frac{8}{10}$ ist. Wie oben bereits gezeigt, suchen wir also ein c, so dass $\Phi(\frac{c\sqrt{n}}{2})$ größer oder gleich $\frac{9}{10}$ ist. Nachdem wir insgesamt 1600 Stichproben vornehmen, approximieren wir c als

$$c \approx \frac{2 \cdot 1{,}28}{\sqrt{1600}} = 0{,}064$$

und erhalten ein Konfidenzintervall von $[\bar{X} - 0{,}064, \bar{X} + 0{,}064]$.

Hausaufgabe 1 (5 Punkte)

Seien x_1 bis x_n unabhängige Stichproben einer Zufallsvariable X. Bestimmen Sie den Maximum-Likelihood-Schätzwert für den Parameter λ unter der Annahme, dass X Poisson-verteilt ist. Bestimmen Sie außerdem den Maximum-Likelihood-Schätzwert für den Parameter p unter der Annahme, dass X geometrisch verteilt ist.

Lösungsvorschlag

Wir betrachten zunächst den Fall, dass X Possion-Verteilt ist. Bei den Stichproben x_1 bis x_n handelt es sich also um natürliche Zahlen einschließlich der 0 und die Dichte einer einzelnen Stichprobe x_i beträgt

$$f_X(x_i) = \frac{\lambda^{x_i}}{e^{\lambda} x_i!}.$$

Des Weiteren ist die Likelihood-Funktion definiert als

$$L(x_1, \dots, x_n; \lambda) = \prod_{i=1}^{n} \frac{\lambda^{x_i}}{e^{\lambda} x_i!} = e^{-n\lambda} \prod_{i=1}^{n} \frac{\lambda^{x_i}}{x_i!}.$$

Wie in der Vorlesung demonstriert, optimieren wir im Folgenden mit Hilfe des Logarithmus der Likelihood-Funktion

$$\ln(L(x_1, \dots, x_n; \lambda)) = -n\lambda + \sum_{i=1}^{n} (x_i \ln(\lambda) - \ln(x_i!)).$$

Leiten wir nun nach λ, ab so erhalten wir den Term

$$\frac{\mathrm{d}}{\mathrm{d}\lambda} - n\lambda + \sum_{i=1}^{n} (x_i \ln(\lambda) - \ln(x_i!)) = -n + \sum_{i=1}^{n} \frac{x_i}{\lambda}$$

mit genau einer Nullstelle für $\lambda = \sum_{i=1}^{n} \frac{1}{n} x_i$. Um zu überprüfen, dass es sich dabei um ein Maximum handelt, leiten wir ein weiteres mal nach λ ab und erhalten

$$\frac{\mathrm{d}}{\mathrm{d}\lambda} - n + \sum_{i=1}^{n} \frac{x_i}{\lambda} = -\sum_{i=1}^{n} \frac{x_i}{\lambda^2}.$$

Offensichtlich ist die zweite Ableitung für keinen Wert von λ positiv und somit muss es sich bei $\sum_{i=1}^{n} \frac{1}{n} x_i$ um ein globales Maximum handeln. Der Maximum-Likelihood-Schätzwert der Poisson-Verteilung entspricht also dem Stichprobenmittel. Untersuchen wir nun den Fall, dass X geometrisch verteilt ist. Die Stichproben x_1 bis x_n sind folglich natürliche Zahlen und mit der Dichte

$$f_X(x_i) = p \cdot (1-p)^{x_i - 1}.$$

Wie im ersten Teil bestimmen wir den Logarithmus der Likelihood-Funktion

$$\ln(L(x_1, \dots, x_n, p)) = \ln\left(\prod_{i=1}^{n} p \cdot (1-p)^{x_i - 1} \right) = n \ln(p) + \left(\sum_{i=1}^{n} x_i - 1 \right) \ln(1-p)$$

und leiten diesen nach p ab

$$\frac{\mathrm{d}}{\mathrm{d}p} n \ln(p) + \left(\sum_{i=1}^{n} x_i - 1 \right) \ln(1-p) = \frac{n}{p} - \frac{\sum_{i=1}^{n} x_i - 1}{1-p}$$

um ein globales Maximum zu finden. Wie bei der Poisson-Verteilung gibt es auch hier lediglich eine Nullstelle, nämlich

$$
\begin{aligned}
&& \frac{n}{p} &= \frac{\sum_{i=1}^{n} x_i - 1}{1-p} \\
\iff && (1-p)n &= \left(\sum_{i=1}^{n} x_i - 1\right) p \\
\iff && n &= \left(n + \sum_{i=1}^{n} x_i - 1\right) p \\
\iff && n &= \left(\sum_{i=1}^{n} x_i\right) p \\
\iff && p &= \frac{n}{\sum_{i=1}^{n} x_i}.
\end{aligned}
$$

Für den Nachweis, dass es sich tatsächlich um Maximum handelt, betrachten wir noch die zweite Ableitung

$$
\frac{\mathrm{d}}{\mathrm{d}p} \frac{n}{p} - \frac{\sum_{i=1}^{n} x_i - 1}{1-p} = -\frac{n}{p^2} - \frac{\sum_{i=1}^{n} x_i - 1}{(1-p)^2}
$$

und stellen fest, dass diese für alle $0 < p < 1$ negativ ist. Das wiederum bedeutet, dass es sich bei $\frac{n}{\sum_{i=1}^{n} x_i}$ in der Tat um den Maximum-Likelihood-Schätzwert der geometrischen Verteilung handelt.

Hausaufgabe 2 (5 Punkte)

Sie befinden sich in einer fremden Stadt, in der es insgesamt n Taxis gibt. Um die Fahrzeuge zu identifizieren, besitzt jedes Taxi eine eindeutigen Nummer von 1 bis n. Während Ihres Aufenthalts fahren Sie zehnmal Taxi, wobei Sie sich die jeweiligen Identifikationsnummern merken. Ihre Wahl der Taxis sei unabhängig und gleichwahrscheinlich. Angenommen die größte Identifikationsnummer, die Sie beobachten, ist 100. Bestimmen Sie eine möglichst kleine obere Schranke für n mit einem Konfidenzniveau von $\frac{9}{10}$.

Lösungsvorschlag

Seien X_1 bis X_{10} die Identifikationsnummern, die Sie während Ihres Aufenthalt beobachten. Des Weiteren sei X das Maximum über alle X_i. Nachdem die Stichprobenvariablen unabhängig sind, ist die Wahrscheinlichkeit, dass X einen Wert kleiner gleich einer natürlichen Zahl $m \leq n$ annimmt, gegeben durch

$$
\Pr[X \leq m] = \Pr[X_1 \leq m, \ldots, X_{10} \leq m] = \prod_{i=1}^{10} \Pr[X_i \leq m] = \prod_{i=1}^{10} \frac{m}{n} = \left(\frac{m}{n}\right)^{10}.
$$

Um eine obere Schranke U zu bestimmen, definieren wir $U = \lceil \delta X \rceil - 1$ und suchen einen möglichst kleinen Wert δ, so dass $\Pr[U \geq n]$ mindestens $\frac{9}{10}$ ist. Durch umformen erhalten wir jetzt einerseits

$$
\Pr[U \geq n] = \Pr[\lceil \delta X \rceil - 1 \geq n] \leq \Pr[\delta X > n] = 1 - \Pr[\delta X \leq n]
$$

und außerdem

$$\Pr[\delta X \leq n] = \Pr\left[X < \frac{n}{\delta}\right] = \Pr\left[X < \left\lceil \frac{n}{\delta} \right\rceil\right] = \left(\frac{\lceil \frac{n}{\delta} \rceil}{n}\right)^{10} \geq \left(\frac{\frac{n}{\delta}}{n}\right)^{10} = \left(\frac{1}{\delta}\right)^{10}.$$

Dies impliziert wiederum, dass δ mindestens $10^{\frac{1}{10}}$ sein muss, da gilt

$$\begin{aligned}
& & \Pr[U \geq n] &\geq \frac{9}{10} \\
&\Longrightarrow & 1 - \Pr[\delta X < n] &\geq \frac{9}{10} \\
&\Longleftrightarrow & \Pr[\delta X < n] &\leq \frac{1}{10} \\
&\Longrightarrow & \left(\frac{1}{\delta}\right)^{10} &\leq \frac{1}{10} \\
&\Longleftrightarrow & \delta &\geq 10^{\frac{1}{10}}
\end{aligned}$$

Nachdem die größte Identifikationsnummer, die Sie beobachtet haben, 100 ist, hat unsere obere Schranke U also einen Wert von

$$\lceil 10^{\frac{1}{10}} 100 \rceil - 1 = 125.$$

Hausaufgabe 3 (5 Punkte)

Sei X eine beliebige Zufallsvariable mit endlichem Erwartungswert und Varianz. In der Vorlesung wurde gezeigt, dass das Stichprobenmittel $\frac{1}{n}\sum_{i=1}^{n} X_i$ ein erwartungstreuer Schätzer für den Erwartungswert $\mathbb{E}[X]$ ist und die Stichprobenvarianz $\frac{1}{n-1}\sum_{i=1}^{n}(X_i - \bar{X})^2$ ein erwartungstreuer Schätzer für die Varianz $\mathrm{Var}[X]$. Im Folgenden möchten wir die beiden Schätzer genauer analysieren.

1. Seien x_1 bis x_n und y_1 bis y_n zwei Reihen unabhängiger Stichproben von X und μ_x bzw. μ_y die durch das Stichprobenmittel geschätzten Erwartungswerte. Des Weiteren bezeichnen wir die auf beiden Reihen basierende Schätzung mit μ_{xy}. Zeigen Sie, dass wenn $\mu_x = \mu_y$ gilt auch $\mu_x = \mu_{xy}$ gelten muss.

2. Beweisen Sie nun, dass diese Eigenschaft auf die Stichprobenvarianz und die entsprechenden Schätzungen σ_x^2, σ_y^2 sowie σ_{xy}^2 nicht notwendigerweise zutrifft.

Lösungsvorschlag

Ist das Stichprobenmittel der x-Reihen mit dem der y-Reihe identische, so gilt

$$\mu_x = \frac{\sum_{i=1}^{n} x_i}{n} = \frac{\sum_{i=1}^{n} y_i}{n} = \mu_y,$$

was wiederum impliziert, dass auch das Stichprobenmittel über beide Reihen den selben Wert hat

$$\mu_{xy} = \frac{\sum_{i=1}^{n} x_i + \sum_{i=1}^{n} y_i}{2n} = \frac{\mu_x}{2} + \frac{\mu_y}{2} = \frac{\mu_x}{2} + \frac{\mu_x}{2} = \mu_x.$$

Als Gegenbeispiel der zweite Teilaufgabe nutzen wir die Stichproben 1, 2 und 3 für die x-Reihe und die Stichproben 3, 4 und 5 für die y-Reihe. Die Stichprobenmittel sind nunmehr $\mu_x = 2$, $\mu_y = 4$ und $\mu_{xy} = 3$. Außerdem berechnen sich die Varianzen der einzelnen Reihen jeweils zu

$$\sigma_x^2 = \frac{(1-2)^2 + (2-2)^2 + (3-2)^2}{3-1} = 1 = \frac{(3-4)^2 + (4-4)^2 + (5-4)^2}{3-1} = \sigma_y^2$$

Für beide Reihen zusammen ergibt sich allerdings ein anderer Wert, nämlich

$$\sigma_{xy}^2 = \frac{(1-3)^2 + (2-3)^2 + 2(3-3)^2 + (4-3)^2 + (5-3)^2)}{5} = 2.$$

was den Beweis vervollständigt.

Hausaufgabe 4 (5 Punkte)

Sie möchten die Anzahl n der Fische im See hinter dem Fakultätsgebäude des Maschinenbaus bestimmen. Hierzu haben Sie sich folgendes Verfahren überlegt. Zunächst fangen Sie eine zufällige Stichprobe von m Fischen und markieren diese geeignet, bevor Sie sie zurück in den See lassen. Nach einiger Zeit fangen Sie eine weitere zufällige Stichprobe von k Fischen. Sei X die Anzahl der bereits markierten Fische in der zweiten Stichprobe. Bestimmen Sie einen Maximum-Likelihood-Schätzer für n basierend auf X.

Hinweis: Betrachten Sie das Verhältnis der Likelihood-Funktion für zwei konsekutive Werte von n, also $\frac{L(x;n)}{L(x;n-1)}$

Lösungsvorschlag

Intuitiv betrachtet sollte der Anteil der markierten Fische in der Stichprobe dem Anteil der markierten Fische im See entsprechen. Wir vermuten demnach, dass der Maximum-Likelihood-Schätzwert für n gleich $\frac{mk}{X}$ ist. Zur Überprüfung dieser Annahme betrachten wir die Likelihood-Funktion von n für eine konkrete Stichprobe x. Da es sich bei X um eine hypergeometrisch verteilte Zufallsvariable handelt, gilt

$$L(x;n) = Pr[X = x] = \frac{\binom{m}{x}\binom{n-m}{k-x}}{\binom{n}{k}}.$$

Um das Optimum der Likelihood-Funktion zu bestimmen, betrachten wir nun das Verhältnis zwischen zwei aufeinanderfolgenden Werten für n. Offensichtlich steigt die Likelihood-Funktion genau dann, wenn $\frac{L(x;n)}{L(x;n-1)}$ größer als 1 ist bzw. $L(x;n)$ größer als $L(x;n-1)$. Wir formen also um

$$\frac{L(x;n)}{L(x;n-1)} = \frac{\binom{n-m}{k-x}\binom{n-1}{k}}{\binom{n-1-m}{k-x}\binom{n}{k}} = \frac{\frac{(n-m)!}{(k-x)!\cdot(n-m-k+x)!} \cdot \frac{(n-1)!}{k!\cdot(n-1-k)!}}{\frac{(n-1-m)!}{(k-x)!\cdot(n-1-m-k+x)!} \cdot \frac{n!}{k!\cdot(n-k)!}} = \frac{(n-m)\cdot(n-k)}{n\cdot(n-m-k+x)},$$

und lösen auf zu

$$\begin{aligned} & (n-m)\cdot(n-k) > n\cdot(n-m-k+x) \\ \iff \quad & n^2 - nm - nk + mk > n^2 - nm - nk + nx \\ \iff \quad & \frac{mk}{x} < n. \end{aligned}$$

Somit haben wir gezeigt, dass die Likelihood-Funktion genau dann steigt, so lange n kleiner als $\frac{mk}{x}$ ist, was wiederum bedeutet, dass wir am Punkt $\frac{mk}{x}$ das Optimum erreichen. Unsere Vermutung hat sich somit bestätigt. Allerdings gilt es zu beachte, dass $\frac{mk}{x}$ nicht notwendigerweise eine ganze Zahl ist. Um einen zulässigen Wert für n zu erhalten, können wir aber zwischen $\lfloor \frac{mk}{x} \rfloor$ und $\lceil \frac{mk}{x} \rceil$ wählen, so dass die Likelihood-Funktion optimiert wird.

$\Phi(x)$	0,00	0,01	0,02	0,03	0,04	0,05	0,06	0,07	0,08	0,09
0,0	0,500	0,504	0,508	0,512	0,516	0,520	0,524	0,528	0,532	0,536
0,1	0,540	0,544	0,548	0,552	0,556	0,560	0,564	0,567	0,571	0,575
0,2	0,579	0,583	0,587	0,591	0,595	0,599	0,603	0,606	0,610	0,614
0,3	0,618	0,622	0,626	0,629	0,633	0,637	0,641	0,644	0,648	0,652
0,4	0,655	0,659	0,663	0,666	0,670	0,674	0,677	0,681	0,684	0,688
0,5	0,691	0,695	0,698	0,702	0,705	0,709	0,712	0,716	0,719	0,722
0,6	0,726	0,729	0,732	0,736	0,739	0,742	0,745	0,749	0,752	0,755
0,7	0,758	0,761	0,764	0,767	0,770	0,773	0,776	0,779	0,782	0,785
0,8	0,788	0,791	0,794	0,797	0,800	0,802	0,805	0,808	0,811	0,813
0,9	0,816	0,819	0,821	0,824	0,826	0,829	0,831	0,834	0,836	0,839
1,0	0,841	0,844	0,846	0,848	0,851	0,853	0,855	0,858	0,860	0,862
1,1	0,864	0,867	0,869	0,871	0,873	0,875	0,877	0,879	0,881	0,883
1,2	0,885	0,887	0,889	0,891	0,893	0,894	0,896	0,898	0,900	0,901
1,3	0,903	0,905	0,907	0,908	0,910	0,911	0,913	0,915	0,916	0,918
1,4	0,919	0,921	0,922	0,924	0,925	0,926	0,928	0,929	0,931	0,932
1,5	0,933	0,934	0,936	0,937	0,938	0,939	0,941	0,942	0,943	0,944
1,6	0,945	0,946	0,947	0,948	0,949	0,951	0,952	0,953	0,954	0,954
1,7	0,955	0,956	0,957	0,958	0,959	0,960	0,961	0,962	0,962	0,963
1,8	0,964	0,965	0,966	0,966	0,967	0,968	0,969	0,969	0,970	0,971
1,9	0,971	0,972	0,973	0,973	0,974	0,974	0,975	0,976	0,976	0,977
2,0	0,977	0,978	0,978	0,979	0,979	0,980	0,980	0,981	0,981	0,982
2,1	0,982	0,983	0,983	0,983	0,984	0,984	0,985	0,985	0,985	0,986
2,2	0,986	0,986	0,987	0,987	0,987	0,988	0,988	0,988	0,989	0,989
2,3	0,989	0,990	0,990	0,990	0,990	0,991	0,991	0,991	0,991	0,992
2,4	0,992	0,992	0,992	0,992	0,993	0,993	0,993	0,993	0,993	0,994
2,5	0,994	0,994	0,994	0,994	0,994	0,995	0,995	0,995	0,995	0,995
2,6	0,995	0,995	0,996	0,996	0,996	0,996	0,996	0,996	0,996	0,996
2,7	0,997	0,997	0,997	0,997	0,997	0,997	0,997	0,997	0,997	0,997
2,8	0,997	0,998	0,998	0,998	0,998	0,998	0,998	0,998	0,998	0,998
2,9	0,998	0,998	0,998	0,998	0,998	0,998	0,998	0,999	0,999	0,999