Kapitel 2

Konstruktion von Schätzern

Wir betrachten im Folgenden die zwei gängigsten Methoden zur Konstruktion von Schätzern - die Maximum Likelihood Methode und die Momentenmethode.

2.1 Maximum Likelihood Schätzer

Sei $X \sim (P_\theta)_{\theta \in \Theta}$, $(X_1, \ldots, X_n)$ eine Zufallsstichprobe von X und $(x_1, \ldots, x_n)$ eine konkrete Realisierung. Wie im letzten Abschnitt verwenden wir dieselbe Notation für den diskreten und den absolut stetigen Fall:

- Falls X diskret mit (vom Parameter θ unabhängigen) Wertebereich $W = \{\alpha_1, \alpha_2, \ldots\}$ ist, schreiben wir
$$f(x; \theta) := P_\theta(X = x)$$
für jedes $x \in W$.
- Falls X absolut stetig ist, bezeichnet $f(x; \theta)$ wie gewohnt die Dichte von P_θ.

Definition 2.1 *Sei $X \sim (P_\theta)_{\theta \in \Theta}$, $(X_1, \ldots, X_n)$ eine Zufallsstichprobe von X und $(x_1, \ldots, x_n)$ eine konkrete Realisierung. Dann heißt die Funktion*

$$L(x_1, \ldots, x_n; \theta) = \prod_{i=1}^{n} f(x_i; \theta) \tag{2.1}$$

Likelihood Funktion (kurz: Likelihood).

Im diskreten Fall entspricht $L(x_1, \ldots, x_n; \theta)$ also genau der Wahrscheinlichkeit

$$P(X_1 = x_1, X_2 = x_2, \ldots, X_n = x_n),$$

im absolut stetigen Fall genau dem Wert der Wahrscheinlichkeitsdichte des Vektors $(X_1, \ldots, X_n)$ an der Stelle $(x_1, \ldots, x_n)$.
Die Grundidee der Maximum Likelihood Methode besteht nun einfach darin, als Schätzer $\hat{\theta}_n$ für θ jenen Wert aus Θ zu verwenden, der $L(x_1, \ldots, x_n; \theta) = \prod_{i=1}^{n} f(x_i; \theta)$ maximiert (falls existent). Im diskreten Fall bedeutet dies also, dass man genau jenen Parameter sucht, für den die beobachtete Realisierung maximale Wahrscheinlichkeit hat.
Bevor wir Maximum Likelihood (kurz: ML) Schätzer allgemein definieren, betrachten wir das folgende einfache Beispiel, das die Nützlichkeit und Güte von ML-Schätzern gut illustriert.

Beispiel 2.2 Sei $X \sim (Bin(1,\theta))_{\theta\in[0,1]}$. Dann erhalten wir für die Likelihood Funktion

$$L(x_1,\ldots,x_n;\theta) = \prod_{i=1}^{n} f(x_i;\theta) = \prod_{i=1}^{n} \theta^{x_i}\,(1-\theta)^{1-x_i} = \theta^{\sum_{i=1}^n x_i}\,(1-\theta)^{n-\sum_{i=1}^n x_i}.$$

Wir unterscheiden drei Fälle:

1. Für $\sum_{i=1}^n x_i = 0$ folgt $L(x_1,\ldots,x_n;\theta) = (1-\theta)^n$, die Likelihood Funktion ist also streng monoton fallend in θ und das Maximum wird in $\theta = 0 = \frac{1}{n}\sum_{i=1}^n x_i$ angenommen.

2. Für $\sum_{i=1}^n x_i = n$ folgt $L(x_1,\ldots,x_n;\theta) = \theta^n$, die Likelihood Funktion ist also streng monoton wachsend in θ und das Maximum wird in $\theta = 1 = \frac{1}{n}\sum_{i=1}^n x_i$ angenommen.

3. $\sum_{i=1}^n x_i \in \{1,2,\ldots,n-1\}$. Nachdem $L(x_1,\ldots,x_n;\theta) \geq 0$ für jedes $\theta \in [0,1]$ und $L(x_1,\ldots,x_n;0) = 0 = L(x_1,\ldots,x_n;1)$ reicht es, $\theta \in (0,1)$ zu betrachten und, zur Vereinfachung der Rechenarbeit, lokale Extrema von $\ln L(x_1,\ldots,x_n;\theta)$ zu berechnen: Wir setzen $k := \sum_{i=1}^n x_i$, erhalten

$$\ln L(x_1,\ldots,x_n;\theta) = k\ln\theta + (n-k)\ln(1-\theta)$$

und damit

$$\begin{aligned}\frac{\partial \ln L(x_1,\ldots,x_n;\theta)}{\partial\theta} &= \frac{k}{\theta} - \frac{n-k}{1-\theta} = \frac{k-\theta n}{\theta(1-\theta)}\\ \frac{\partial^2 \ln L(x_1,\ldots,x_n;\theta)}{\partial\theta^2} &= -\frac{k}{\theta^2} - \frac{n-k}{(1-\theta)^2} < 0.\end{aligned}$$

Damit ist $\ln L(x_1,\ldots,x_n;\theta)$ genau dann maximal wenn $\theta = \frac{k}{n} = \frac{1}{n}\sum_{i=1}^n x_i$.

Alle drei Fälle liefern also dasselbe Resultat

$$\hat{\theta}_n = \frac{1}{n}\sum_{i=1}^{n} X_i = \overline{X}_n,$$

i.e. der durch Maximierung des Likelihood erhaltenen Schätzer ist genau $\overline{X}_n$. Nachdem $\overline{X}_n$ laut Beispiel 10.15 effizient für θ ist, liefert also die Maximierung des Likelihood einen effizienten und noch dazu stark konsistenten Schätzer.

Definition 2.3 *Sei $X \sim (P_\theta)_{\theta\in\Theta}$ und $(X_1,\ldots,X_n)$ eine Zufallsstichprobe von X. Dann heißt ein Schätzer $\hat{\theta}_n$ von θ* Maximum Likelihood Schätzer (MLE) *genau dann wenn*

$$L(X_1,\ldots,X_n;\hat{\theta}_n) = \sup_{\theta\in\Theta} L(X_1,\ldots,X_n;\theta),$$

i.e. wenn er den Likelihood maximiert.

Bemerkung 2.4 Maximum Likelihood Schätzer müssen (im Allgemeinen) weder existieren noch eindeutig sein.

Beispiel 2.5 Wir betrachten den Fall $X \sim \mathcal{N}(\mu, \sigma^2)$ mit $\theta = (\theta_1, \theta_2) = (\mu, \sigma^2) \in \Theta = \mathbb{R} \times (0, \infty)$ und berechnen den MLE für θ. Die Likelihood Funktion hat folgende Form:

$$L(x_1, \ldots, x_n; \theta_1, \theta_2) = \prod_{i=1}^{n} \frac{1}{\sqrt{2\pi\theta_2}} e^{-\frac{(x_i-\theta_1)^2}{2\theta_2}} = (2\pi\theta_2)^{-n/2} e^{-\frac{1}{2\theta_2}\sum_{i=1}^{n}(x_i-\theta_1)^2}$$

Wir können zur Berechnung des Maximums die partiellen Ableitungen von $\ln L(x_1, \ldots, x_n; \theta_1, \theta_2)$ nach θ_1 und θ_2 betrachten und Null setzen:

$$\ln L(x_1, \ldots, x_n; \theta_1, \theta_2) = -\frac{n}{2} \ln(2\pi\theta_2) - \frac{1}{2\theta_2} \sum_{i=1}^{n} (x_i - \theta_1)^2$$

$$\begin{aligned}
\frac{\partial \ln L(x_1, \ldots, x_n; \theta_1, \theta_2)}{\partial \theta_1} &= \frac{1}{\theta_2} \sum_{i=1}^{n} (x_i - \theta_1) \stackrel{!}{=} 0 \\
\frac{\partial \ln L(x_1, \ldots, x_n; \theta_1, \theta_2)}{\partial \theta_2} &= -\frac{n}{2\theta_2} + \frac{1}{2\theta_2^2} \sum_{i=1}^{n} (x_i - \theta_1)^2 \stackrel{!}{=} 0
\end{aligned}$$

Als Lösung erhalten wir daher sofort

$$\hat{\theta}_n = \left((\hat{\theta_1})_n, (\hat{\theta_2})_n\right) = \left(\overline{X}_n, \frac{1}{n} \sum_{i=1}^{n} (X_i - \overline{X}_n)^2\right).$$

Aus den vorhergehenden Kapiteln wissen wir, dass sowohl $(\hat{\theta_1})_n$ als auch $(\hat{\theta_2})_n$ stark konsistente Schätzer von $\theta_1 = \mu$ bzw. $\theta_2 = \sigma^2$ sind. Weiters ist $(\hat{\theta_1})_n$ erwartungstreu, $(\hat{\theta_2})_n$ asymptotisch erwartungstreu.

Bemerkung 2.6 Wie vorhin schon erwähnt entspricht der MLE im diskreten Fall jenem Parameter θ für den die beobachtete Realisierung maximale Wahrscheinlichkeit hat. Obwohl diese Interpretation offensichtlich nicht direkt auf den absolut stetigen Fall übertragen werden kann (jede Realisierung hat Wahrscheinlichkeit Null) kann das 'Funktionieren' der MLE Methode auch in diesem Fall wiefolgt *heuristisch* begründet werden: Angenommen $X \sim P_{\theta_0}$ für ein $\theta_0 \in \Theta$ und $(X_1, \ldots, X_n)$ ist eine Zufallsstichprobe von X. Unter der Voraussetzung, dass $\ln(f(X_1; \theta))$ integrierbar ist liefert das SLLN

$$\frac{1}{n} \ln L(X_1, \ldots, X_n; \theta) = \frac{1}{n} \sum_{i=1}^{n} \ln(f(X_i; \theta) \xrightarrow{[P]} \mathbb{E}_{\theta_0}(\ln(f(X_1; \theta)) = \int_{\mathbb{R}} \ln(f(x; \theta)) f(x, \theta_0) dx.$$

Unter Verwendung der Ungleichung $\ln x \leq x - 1$ folgt weiter

$$\begin{aligned}
\mathbb{E}_{\theta_0}(\ln(f(X_1; \theta)) - \mathbb{E}_{\theta_0}(\ln(f(X_1; \theta_0)) &= \int_{\mathbb{R}} \ln(f(x; \theta)) f(x; \theta_0) dx - \int_{\mathbb{R}} \ln(f(x; \theta_0)) f(x; \theta_0) dx \\
&= \int_{\mathbb{R}} \ln\left(\frac{f(x; \theta)}{f(x; \theta_0)}\right) f(x; \theta_0) dx \\
&\leq \int_{\mathbb{R}} \left(\frac{f(x; \theta)}{f(x; \theta_0)} - 1\right) f(x; \theta_0) dx = 1 - 1 = 0,
\end{aligned}$$

und damit

$$\frac{1}{n} \ln L(X_1, \ldots, X_n; \theta) \doteq \mathbb{E}_{\theta_0}(\ln(f(X_1; \theta)) \leq \mathbb{E}_{\theta_0}(\ln(f(X_1; \theta_0)).$$

Mit anderen Worten, $L(X_1, \ldots, X_n; \theta)$ sollte asymptotisch für $\theta = \theta_0$ maximal werden.

Allgemein haben Maximum-Likelihood Schätzer für den Fall, dass $f(x;\theta)$ gewisse, nicht allzu restriktive Regularitätsbedingungen erfüllt (im Wesentlichen Glattheit von $f(x;\theta)$, Kompaktheit des Parameterraums Θ und Eindeutigkeit des MLE) gute Eigenschaften. Nachdem wir an dieser Stelle nicht näher auf diese Bedingungen eingehen können, zählen wir die wesentlichsten 'guten' asymptotischen Eigenschaften nur ohne Beweis und genaue Formulierung der Voraussetzen im folgenden Satz auf.

Satz 2.7 *Sei $X \sim (P_\theta)_{\theta\in\Theta}$ mit $\Theta \subseteq \mathbb{R}$ und $\hat{\theta}_n$ Maximum Likelihood Schätzer von θ. Unter gewissen Regularitätsvoraussetzung hat $\hat{\theta}_n$ die folgenden Eigenschaften:*

1. *$\hat{\theta}_n$ ist konsistent für θ, i.e. für jedes feste $\theta_0 \in \Theta$ gilt: falls $X \sim P_{\theta_0}$ dann folgt $\hat{\theta}_n \xrightarrow{P} \theta_0$.*

2. *$\hat{\theta}_n$ ist invariant, i.e. für eine messbare Funktion g ist $g \circ \hat{\theta}_n$ Maximum Likelihood Schätzer von $g(\theta)$.*

3. *$\hat{\theta}_n$ ist asymptotisch normal, i.e. für jedes feste $\theta_0 \in \Theta$ gilt: falls $X \sim P_{\theta_0}$ dann folgt*

$$\sqrt{n}(\hat{\theta}_n - \theta_0) \xrightarrow{w} Z$$

wobei $Z \sim \mathcal{N}(0, \frac{1}{I(\theta_0)})$.

Beispiel 2.8 Wir kehren nochmals zurück zur in Beispiel 2.2 beschriebenen Situation. Wie schon erwähnt ist $\hat{\theta}_n = \overline{X}_n$ sogar stark konsistent und effizient, es gilt also etwas mehr als Aussage eins in Satz 2.7. Falls $X \sim Bin(1, \theta_0)$, dann impliziert das CLT (Satz 8.15), dass

$$Z_n = \sqrt{n}\frac{\overline{X}_n - \theta_0}{\sqrt{\theta_0(1-\theta_0)}} \xrightarrow{w} Z$$

wobei $Z \sim \mathcal{N}(0,1)$. Wegen $I(\theta_0) = \frac{1}{\theta_0(1-\theta_0)}$ folgt daraus sofort die dritte Aussage in Satz 2.7.

2.2 Momentenschätzer

Wir beginnen mit zwei einfachen Beispielen, die die, der Momentenmethode zugrundeliegende Idee veranschaulichen:

Beispiel 2.9 Sei $X \sim (\mathcal{E}(\theta))_{\theta \in (0,\infty)}$. Für Zufallsstichproben $X_1, X_2, \ldots, X_n$ von X impliziert das SLLN

$$\lim_{n\to\infty} \frac{1}{n}\sum_{i=1}^{n} X_i \xrightarrow{[P]} \mathbb{E}_\theta(X) = \frac{1}{\theta} =: g(\theta).$$

Nachdem $g : (0,\infty) \to (0,\infty)$ bijektiv und stetig ist, ist

$$\hat{\theta}_n := g^{-1}\Big(\frac{1}{n}\sum_{i=1}^{n} X_i\Big) = \frac{1}{\overline{X}_n}$$

ein stark konsistenter Schätzer von θ (vergleiche mit dem MLE für θ in den Übungen).

Beispiel 2.10 Sei X absolut stetig mit Dichte $f_\theta(x) = (\theta+1)x^\theta\, \mathbf{1}_{(0,1)}(x)$ für ein $\theta > 0$. Wir versuchen analog zum vorigen Beispiel, einen 'guten' Schätzer für den Parameter θ zu finden. Mit dem SLLN im Hinterkopf berechnen wir als ersten Schritt $\mathbb{E}(X)$ und erhalten

$$\mathbb{E}_\theta(X) = \int_{\mathbb{R}} x f_\theta(x)\, d\lambda(x) = (\theta+1)\int_{(0,1)} x^{\theta+1} d\lambda(x) = \frac{\theta+1}{\theta+2} =: g(\theta).$$

Für Zufallsstichproben $X_1, X_2, \ldots, X_n$ von X impliziert das SLLN

$$\lim_{n\to\infty} \frac{1}{n}\sum_{i=1}^{n} X_i \xrightarrow{[P]} \mathbb{E}_\theta(X) = \frac{\theta+1}{\theta+2} = g(\theta).$$

Nachdem $g : (-1,\infty) \to (0,1)$ bijektiv und stetig ist, folgt sofort nach Satz 9.8, dass

$$\hat{\theta}_n := g^{-1}(\overline{X}_n) = \frac{1}{1-\overline{X}_n} - 2$$

ein stark konsistenter Schätzer von θ ist. Weitere Eigenschaften von $\hat{\theta}_n$ werden in den Übungen untersucht. Abbildung 2.10 zeigt ein Histogramm der Schätzwerte $\hat{\theta}_n$ für den Fall $\theta = 3$, $n = 100$ ($n = 1000$) und $R = 10000$ Wiederholungen.

Für $X \sim (P_\theta)_{\theta\in\Theta}$ mit $\Theta \subseteq \mathbb{R}^d$ kann analog vorgegangen werden: Angenommen X^d ist integrierbar und es existieren Funktionen $g_1, g_2, \ldots, g_d$ sodass $\mathbb{E}_\theta(X^k) = g_k(\theta_1, \theta_2, \ldots, \theta_d) = g_k(\theta)$ für jedes $k \in \{1, \ldots, d\}$. Dann impliziert das SLLN

$$\begin{aligned}
\frac{1}{n}\sum_{i=1}^{n} X_i &\xrightarrow{[P]} g_1(\theta_1, \ldots, \theta_d)\\
\frac{1}{n}\sum_{i=1}^{n} X_i^2 &\xrightarrow{[P]} g_2(\theta_1, \ldots, \theta_d)\\
&\ \ \vdots \qquad\qquad \vdots\\
\frac{1}{n}\sum_{i=1}^{n} X_i^d &\xrightarrow{[P]} g_d(\theta_1, \ldots, \theta_d)
\end{aligned}$$

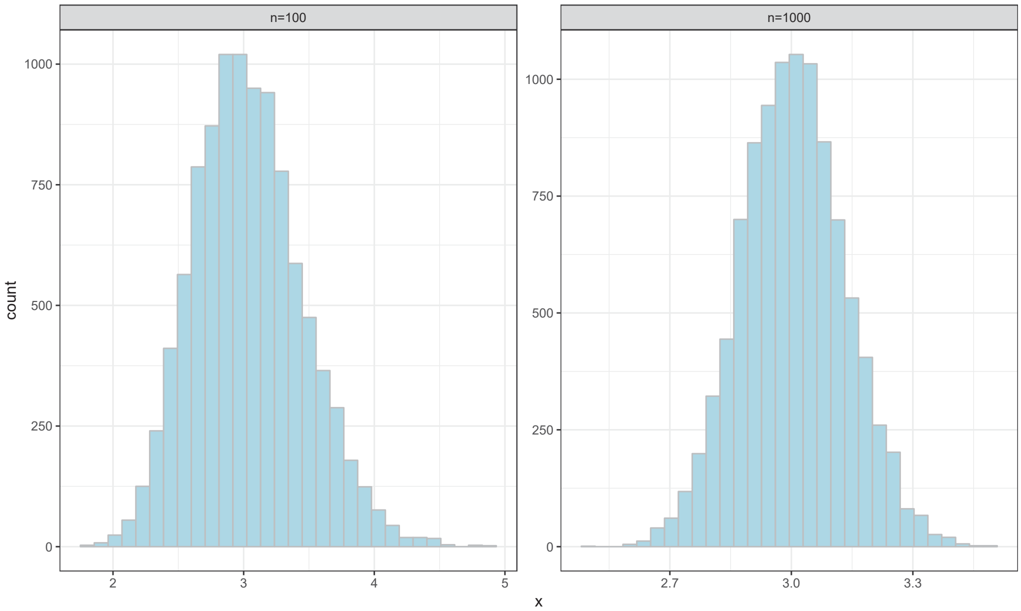

Abbildung 2.1: Histogramm der Schätzwerte $\hat{\theta}_n$ aus Beispiel 2.10 für den Fall $\theta = 3$, $n = 100$ ($n = 1000$) und $R = 10000$ Wiederholungen

und wir erhalten einen Schätzer $\hat{\theta}_n = ((\hat{\theta}_1)_n, (\hat{\theta}_2)_n, \ldots, (\hat{\theta}_d)_n)$ für θ falls das d-dimensionale Gleichungssystem

$$\left.\begin{aligned} m_1 &:= \frac{1}{n}\sum_{i=1}^n X_i = g(\theta_1, \ldots, \theta_d) \\ m_2 &:= \frac{1}{n}\sum_{i=1}^n X_i^2 = g_2(\theta_1, \ldots, \theta_d) \\ &\qquad\vdots \qquad\quad \vdots \\ m_d &:= \frac{1}{n}\sum_{i=1}^n X_i^d = g_d(\theta_1, \ldots, \theta_d) \end{aligned}\right\} =: \mathbf{g}(\theta) \tag{2.2}$$

eindeutig nach θ auflösbar ist. Der resultierende Schätzer $\hat{\theta}_n = \mathbf{g}^{-1}(m_1, \ldots, m_d)$ heißt *Momentenschätzer von θ*. $\hat{\theta}_n = \mathbf{g}^{-1}(m_1, \ldots, m_d)$ ist automatisch stark konsistent wenn $\mathbf{g}$ stetig ist.

Beispiel 2.11 Sei $X \sim \mathcal{N}(\theta_1, \theta_2)_{(\theta_1,\theta_2)\in\mathbb{R}\times(0,\infty)}$. Wir berechnen den Momentenschätzer von $\theta = (\theta_1, \theta_2)$: Wegen $\mathbb{E}_\theta(X) = \theta_1$ und $\mathbb{E}_\theta(X^2) = \mathbb{V}_\theta(X) + \mathbb{E}_\theta(X)^2 = \theta_2 + \theta_1^2$ nimmt das Gleichungssystem (2.2) die folgende einfache Form an:

$$\left.\begin{aligned} m_1 &:= \frac{1}{n}\sum_{i=1}^n X_i = \theta_1 = g_1(\theta_1, \theta_2) \\ m_2 &:= \frac{1}{n}\sum_{i=1}^n X_i^2 = \theta_2 + \theta_1^2 = g_2(\theta_1, \theta_2) \end{aligned}\right\} =: \mathbf{g}(\theta) \tag{2.3}$$

und wir erhalten als Momentenschätzer sofort

$$\hat{\theta}_n = \Big(\overline{X}_n, \frac{1}{n}\sum_{i=1}^n X_i^2 - (\overline{X}_n)^2\Big) = \Big(\overline{X}_n, \frac{1}{n}\sum_{i=1}^n (X_i - \overline{X}_n)^2\Big),$$

also genau den im vorigen Abschnitt erhaltenen MLE (der asmyptotisch erwartungstreu und stark konsistent ist).

Genauso wie Maximum Likelihood Schätzer haben auch Momentenschätzer unter nicht allzu restriktiven Regularitätsbedingungen gute asymptotische Eigenschaften - wie im vorigen Abschnitt formulieren wir nur das eindimensionale Resultat ohne genaue Angabe hinreichender Bedingungen; analoge Eigenschaften gelten auch im allgemeinen multivariaten Fall $\Theta \subseteq \mathbb{R}^d$:

Satz 2.12 *Sei $X \sim (P_\theta)_{\theta\in\Theta}$ mit $\Theta \subseteq \mathbb{R}$ und $\hat{\theta}_n$ der Momentenschätzer von θ. Weiters gelte $g(\theta) = \mathbb{E}_\theta(X)$ für alle $\theta \in \Theta$. Unter gewissen Regularitätsvoraussetzung hat $\hat{\theta}_n$ die folgenden Eigenschaften:*

1. *$\hat{\theta}_n$ ist konsistent für θ, i.e. für jedes feste $\theta_0 \in \Theta$ gilt: falls $X \sim P_{\theta_0}$ dann folgt $\hat{\theta}_n \xrightarrow{P} \theta_0$.*
2. *$\hat{\theta}_n$ ist asymptotisch normal, i.e. für jedes feste $\theta_0 \in \Theta$ gilt: falls $X \sim P_{\theta_0}$ dann folgt*

$$\sqrt{n}(\hat{\theta}_n - \theta_0) \xrightarrow{w} Z$$

wobei $Z \sim \mathcal{N}\Big(0, \frac{\mathbb{V}_{\theta_0}(X)}{(g'(\theta_0))^2}\Big)$.

Beispiel 2.13 Wir betrachten wieder die in Beispiel 2.9 behandelte Situation $X \sim (\mathcal{E}(\lambda))_{\lambda\in(0,\infty)}$. Die Existenz des Schätzers $\hat{\theta}_n = 1/\overline{X}_n$ ist klar, starke Konsistenz ebenfalls. Auch die im vorhergehenden Satz behauptete asymptotische Normalität lässt sich mit Hilfe des CLT und der sogenannten δ-*Methode*, die wir später noch allgemein kennenlernen werden, und hier nur als im univariaten Fall formulieren, beweisen.

Satz 2.14 (Spezialfall δ Methode) *Die Abbildung $\Phi : U \subseteq \mathbb{R} \to \mathbb{R}$ sei differenzierbar im Punkt $\theta \in U$ und erfülle $\Phi'(\theta) \neq 0$. Die Zufallsvariable $T, T_1, T_2, \ldots$ seien U-wertig. Dann impliziert $\sqrt{n}(T_n - \theta) \xrightarrow{w} T \sim \mathcal{N}(0, \sigma^2)$ für $n \to \infty$ die Gültigkeit von*

$$\sqrt{n}(\Phi(T_n) - \Phi(\theta)) \xrightarrow{w} \mathcal{N}(0, \sigma^2(\Phi'(\theta)^2))$$

für $n \to \infty$.

Beispiel 2.15 (Fortsetzung von Beispiel 2.13) Aus dem CLT wissen wir $Z_n \xrightarrow{w} Z \sim \mathcal{N}(0,1)$ wobei

$$Z_n = \sqrt{n}\,\frac{\overline{X}_n - \frac{1}{\lambda}}{\frac{1}{\lambda}}.$$

Insbesondere gilt also

$$\sqrt{n}\left(\overline{X}_n - \frac{1}{\lambda}\right) \xrightarrow{w} Z \sim \mathcal{N}\Big(0, \frac{1}{\lambda^2}\Big).$$

Wir setzen $\theta := \frac{1}{\lambda}$ sowie $\Phi(\theta) = \frac{1}{\theta}$ und erhalten als direkte Folgerung von Satz 2.14 sofort

$$\sqrt{n}\left(\frac{1}{\overline{X}_n} - \lambda\right) \xrightarrow{w} Z \sim \mathcal{N}\Big(0, \frac{1}{\lambda^2}\lambda^4\Big) = \mathcal{N}(0, \lambda^2).$$

Für grosses n erhalten wir damit $\hat{\theta}_n - \lambda \approx \mathcal{N}\Big(0, \frac{\lambda^2}{n}\Big)$, also die gewünschte asymptotische Normalität.

Beispiel 2.16 Für den Fall $X \sim Bin(1, \theta)_{\theta \in [0,1]}$ ist der Momentenschätzer offensichtlich gegeben durch $\hat{\theta}_n = \overline{X}_n$, er stimmt also mit dem MLE überein und ist stark konsistent. Wir haben schon in Beispiel 2.8 gezeigt, dass für $X \sim Bin(1, \theta_0)$

$$Z_n = \sqrt{n} \frac{\hat{\theta}_n - \theta_0}{\sqrt{\theta_0(1-\theta_0)}} \xrightarrow{w} Z$$

gilt, wobei $Z \sim \mathcal{N}(0,1)$. Nachdem in diesem Fall $g' = 1$ und $\mathbb{V}_{\theta_0}(X) = \theta_0(1-\theta_0)$ erhalten wir also genau die zweite, im vorhergehenden Satz behauptete Eigenschaft.