Maximum-Likelihood-Prinzip

Beim Maximum-Likelihood-Ansatz unterscheidet man grundsätzlich zwei Situationen:

A. Bevor die Stichprobe tatsächlich gezogen wird, ist nur die Verteilung der möglichen Stichprobenergebnisse in Abhängigkeit vom Parameterwert θ bekannt:

$$f_{X_1,X_2,\ldots,X_n}(x_1,x_2,\ldots,x_n;\theta)=\prod_{i=1}^{n} f_X(x_i;\theta)$$

D.h. die gemeinsame Verteilung hängt von $x_1,\ldots,x_n$ und dem unbekannten θ ab.

B. Nachdem die Stichprobe gezogen wurde, ist der Wert der Verteilung des konkreten Stichprobenergebnisses in Abhängigkeit vom Parameterwert θ bekannt:

$$L(\theta;x_1,x_2,\ldots,x_n)=\prod_{i=1}^{n} f_X(x_i;\theta)$$

D.h. die gemeinsame Verteilung hängt nur noch von dem unbekannten θ ab, da $x_1,\ldots,x_n$ bekannt sind.

Likelihoodfunktion

- Die Funktion

$$L(\theta)=L(\theta;x_1,\ldots,x_n)$$

heißt Likelihoodfunktion.

- Diskrete Zufallsvariable: Likelihhoodfunktion = Gemeinsame Wahrscheinlichkeitsfunktion.
- Stetige Zufallsvariable: Likelihoodfunktion = Gemeinsame Dichtefunktion (keine Wahrscheinlichkeit).

Likelihoodfunktion für eine Bernoulli-verteilte Zufallsvariable I

- Sei X die Zufallsvariable, die den Wert 1 annimmt, wenn beim Münzwurf "Kopf" oben liegt. Der unbekannte Parameter ist die Wahrscheinlichkeit p, dass "Kopf" oben liegt.
- Bevor eine Stichprobe gezogen wird, d.h. ein 10-maliger Münzwurf vorgenommen wird, kennt man in Abhängigkeit von dem unbekannten p die gemeinsame Verteilung der Stichprobenergebnisse:

$$P(X_1=x_1,\ldots,X_n=x_n;p)=\prod_{i=1}^{n} p^{x_i}(1-p)^{1-x_i}.$$

- Nach der Stichprobenziehung liegt z.B. ein Stichprobenbefund

x_1	x_2	x_3	x_4	x_5	x_6	x_7	x_8	x_9	x_{10}
1	0	0	1	1	0	0	1	1	1

vor.

Likelihoodfunktion für eine Bernoulli-verteilte Zufallsvariable II

- Die Wahrscheinlichkeit dieses Stichprobenergebnisses in Abhängigkeit von p ist

$$L(p)=p^6(1-p)^4.$$

- Wahl von p: Wähle p so, dass die Wahrscheinlichkeit am größten ist, dass der konkrete Stichprobenbefund aus einer Grundgesamtheit mit diesem Parameterwert p stammt.
- D.h.: Maximiere die Likelihoodfunktion bez. p.
- Vereinfachung der Maximierung: Da nur die Stelle des Maximums und nicht der Wert der Zielfunktion an der Stelle des Maximums interessiert, kann auch die logarithmierte Likelihoodfunktion

$$\ln L(p)=\ln L(p;x_1,\ldots,x_n)=6\ln p+4\ln(1-p)$$

maximiert werden.
$L(p)$ und $\ln L(p)$ besitzen das Maximum an derselben Stelle p^*.

Likelihoodfunktion für eine Bernoulli-verteilte Zufallsvariable III

- Ergebnis der Maximierung:

$$\frac{\partial \ln L(p)}{\partial p} = \frac{6}{p} - \frac{4}{1-p} \stackrel{!}{=} 0$$

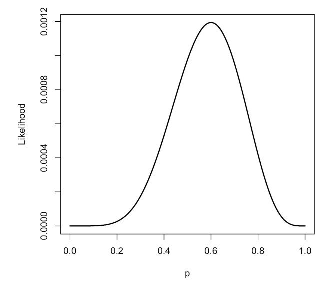
liefert $\hat{p}_{ML} = 6/10 = 0.6$.

- D.h. am wahrscheinlichsten ist es, dass der Stichprobenbefund aus einer Grundgesamtheit mit $p = 0.6$ stammt, also mit 60%-iger Wahrscheinlichkeit "Kopf" oben liegt.
- Wie sähe das Ergebnis aus für eine 100malige Wiederholung des Münzwurf, wobei 55mal "Kopf" oben liegt?

Likelihoodfunktion in **R** I

- Plotten der Likelihoodfunktion für $n = 10$ und $\sum_{i=1}^{10} = 6$:

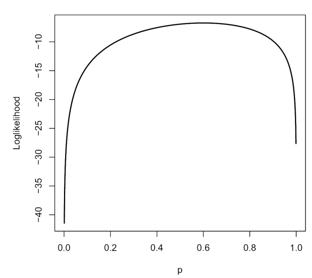
```
>p=seq(0,1,0.001)
>l=p^6*(1-p)^4
>plot(p,l,"l",lwd=2)
```



Likelihoodfunktion in **R** II

- Plotten der logarithmierten Likelihoodfunktion für $n = 10$ und $\sum_{i=1}^{10} = 6$:

```
>p=seq(0,1,0.001)
>l=p^6*(1-p)^4
>ll=log(l)
>plot(p,ll,"l",lwd=2)
```



Maximum-Likelihood-Schätzung für eine Bernoulli-verteilte Zufallsvariable I

- Es soll eine allgemeine Schätzformel angegeben werden, die nicht nur für den Stichprobenbefund $n = 10$ und $\sum_{i=1}^{10} x_i = 6$ einen Schätzwert liefert.
- Likelihoodfunktion für ein beliebiges Stichprobenergebnis $x_1, x_2, \ldots, x_n$:

$$L(p) = \prod_{i=1}^{n} p^{x_i}(1-p)^{1-x_i}$$

- Logarithmierte Likelihoodfunktion für ein beliebiges Stichprobenergebnis $x_1, x_2, \ldots, x_n$:

$$\ln L(p) = \sum_{i=1}^{n} (x_i \ln p + (1 - x_i)\ln(1-p)).$$

Maximum-Likelihood-Schätzung für eine Bernoulli-verteilte Zufallsvariable II

- Notwendige Maximumsbedingung:

$$\frac{\partial \ln L(p)}{\partial p} = \sum_{i=1}^{n} \left(\frac{x_i}{p} - \frac{1 - x_i}{1 - p} \right) \stackrel{!}{=} 0.$$

- Durchmultiplizieren mit $p(1-p) > 0$ und Auflösen nach p ergibt den ML-Schätzer:

$$\hat{p}_{ML} = \frac{1}{n} \sum_{i=1}^{n} X_i$$

- D.h. die unbekannte Wahrscheinlichkeit, mit der "Kopf" beim Münzwurf oben liegen bleibt, wird durch den Anteil der "Köpfe" in der Stichprobe geschätzt.
- Hinreichende Maximumsbedingung:

$$\frac{\partial^2 \ln L(p)}{\partial p^2} = \sum_{i=1}^{n} \left(-\frac{x_i}{p^2} - \frac{1 - x_i}{(1 - p)^2} \right) < 0$$

für alle $p \in (0, 1)$. Somit besitzt die logarithmierte Likelihoodfunktion ein gleichmäßiges (konkaves) Krümmungsverhalten, so dass ein globales Maximum vorliegt.