Seminararbeit: Information einer Stichprobe

Autor: Yingying Zhou

Leitung: PD. Dr. Dr. Christina Schneider

Betreuer: Paul Fink

23.Januar 2015

Inhaltsverzeichnis

1 Einleitung

Der folgende Artikel ist eine Zusammenfassung von Dr. B. Rügers Lehrbuch Test- und Schätztheorie,Band I:Grundlagen(1999): aus Kapitel 1.32-1.33. Mit dem Begriffe der Stichprobe fangen wir mit diesem Artikel an. Der Stichprobe verbindet die folgende Vorstellung: Eine Teilmenge von n Elementen der Grundgesamtheit, wird bezüglich eines interessierenden Merkmals besucht. Das Merkmal wird als Zufallsgröße X definiert, und der spezielle Merkmalswert, den ein Element der Grundgesamtheit aufweist, ist eine Realisierung x von X. Die Zielsetzung der Stichprobe ist die Information über die Wahrscheinlichkeitsverteilung der Zufallsgröße zu gewinnen. Informationen über die unbekannte Verteilung ist eine Zufallsstichprobe. Die Zufallsgröße X der Stichprobe wird n-mal beobachtet. So ergebt eine zufällige Vektor $\{X_1, X_2, \ldots, X_n\}$ mit unabhängigen, identisch wie X verteilten Komponenten X_i. Nach i-ten Ablauf der Versuchung entsteht eine Realisierung x_i von X_i: $X_i = x_i$. Die Vektor von der Realisierung ist $\{x_1, x_2, \ldots, x_n\}$. Information einer Stichprobe ist nämlich die Information einer Beobachtung einer Zufallsgröße über die unbekannte Verteilung. In die folgende Artikel besteht immer eine Verteilungsannahme $\mathfrak{B} = \{P_\theta : \theta \in \Theta\}$ über die Stichprobe X mit einem unbekannten Parameter θ.[B.Rüger (1999)] (Seite 91) Die Verteilungsannahme $\mathfrak{B} = \{f(x;\theta) : \theta \in \Theta\}$, die voraussetzenden Regularitätseigenschaften besitzt, nennen wir Fisher-reguläre Verteilungsfamilie. Durch der Derivation und Logarithmus der Dichtfunktion von Fisher-reguläre Verteilungsannahme bekommen wir die sogenannte Score-Funktion $K(\theta; X)$ der Stichprobe X als die Grundlage zur Definition der Fisher-Information $I_X(\theta)$. Die Fisher-Information ist Varianz von Score oder Erwartungswet der beobachtete Information $I(\theta; x)$. Sie ist eine Schätzung der Größe von der Beobachtungen der Stichprobe X über den Parameter θ. Eine suffiziente Statistik enthlt nur die Informationen der Stichprobe über θ. Die Fisher-Information $I_{T(X)}(\theta)$ von $T = T(X)$ über θ besitzen überstimmende Score-Funktion mit derjenigen von X. Als Folgerung ergibt sich: $I_{T(X)}(\theta) = I_X(\theta)$ für alle θ. Information der Verteilungsannahme mit einem Parametervektor $\theta = (\theta_1, \ldots, \theta_r)$ nennen wir Fishersche Informationsmatrix. Das Grundprinzip der Fishersche Informationsmatrix ist analog zu Fisher-Information.

2 Information einer Stichprobe

Information einer Stichprobe enthalten wir die Informationne über eine unbekannten Paramter θ aus einem zugrundliegenden Verteilungsfamilie $\mathfrak{B}$ über die Stichprobe X. Das Problem ist, wie viele Informationen kann eine Stichprobe über die unbekannten Parameter θ abschätzen. In diesem Kapitel betrifft eine Verteilungsannahme $\mathfrak{B} = \{P_\theta : \theta \in \Theta\}$ über die Stichprobe X.

2.1 Regularität der Verteilungsfamilie

Um die Information einer Stichprobe zu erfassen, stets eine Verteilungsannahme $\mathfrak{B} = \{P_\theta : \theta \in \Theta\}$ über die Stichprobe X mit einem unbekannte Parameter θ. Wir fangen erstens mit der Eigenschaften der Verteilungsannahme $\mathfrak{B}$ einer Stichprobe an.

Eine Verteilungsannahme $\mathfrak{B} = \{P_\theta : \theta \in \Theta\}$ wird mit einem reellem Parameter θ beschränkt, somit $\Theta \subset \mathbb{R}$. Die sogenante Fisher-reguläre Verteilungsfamilie $\mathfrak{B}$ soll mit den folgenden Regularitätseigenschaften [B.Rüger (1999)](Seite 92) anpassen.

1) $\mathfrak{B}$ ist durch ein (σ-finites) Maß v dominiert, also darstellbar als $\mathfrak{B} = \{f(x;\theta) : \theta \in \Theta\}$.
2) Θ ist ein offenes Intevall.
3) Die Verteilungen von $\mathfrak{B}$ haben alle denselben Träger, so daß auch $C := \{x : f(x;\theta) > 0\}$ unabhängig von θ ist.
4) Die Dichten $f(x;\theta)$ sind für jedes $x \in C$ zweimal nach θ differenzierbar, die betreffenden Ableitungen sind integrierbar über x bzgl. v, wobei die betreffenden Differentiationen und Integrationen vertausbar sind, d.h:

$$\int \frac{d}{d\theta} f(x;\theta) dv = \frac{d}{d\theta} \int f(x;\theta) dv = 0 \tag{2.1}$$

$$\int \frac{d^2}{d\theta^2} f(x;\theta) dv = \frac{d}{d\theta} \int \frac{d}{d\theta} f(x;\theta) dv = 0 \tag{2.2}$$

2.2 Fisher-Information

2.2.1 Score-Funktion

Score-Funktion $K(\theta; X)$ ist die wichtige Kenntnis in diesem Kapitel. Sie wird als die Grundlage der Definition von Fisher-Information angesehen. Die Score-Funktion ist eine Differentiation von einer logarithmische Dichtfunktion über den Parameter θ.
Die Score-Funktion einer Fisher-reguläre Verteilungsannahme $\mathfrak{B} = \{f(x;\theta) : \theta \in \Theta\}$ über einer Stichprobe $X = (X_1, \ldots, X_n)$ für jedes $x \in C$ lautet:

$$K(\theta; x) := \frac{d}{d\theta} log f(x;\theta) = \frac{d}{d\theta} f(x;\theta)/f(x;\theta) \tag{2.3}$$

Wir nennen $K(\theta; x)$ die Score-Funktion der Beobachtung x und $K(\theta; X)$ die Score-Funktion der Stichprobe X. Sie zeigt an, wie empfindlich die Dichte $f(x;\theta)$ von der Parameter θ abhängt.
Zunächst zeigen wir die Eigenschaften der Score-Funktion und beweisen wir die noch.

$$E_\theta K(\theta; X) = 0 \tag{2.4}$$

Beweis: $E_\theta K(\theta; X) = \int K(\theta; x) f(x;\theta) dv \overset{2.1}{=} \int \frac{d}{d\theta} f(x;\theta) dv$

$$Var_\theta K(\theta; X) = -E_\theta \frac{d}{d\theta} K(\theta; X) \tag{2.5}$$

Beweis: $Var_\theta K(\theta; X) = E_\theta [K(\theta; X)]^2 = \int \frac{d^2}{d\theta^2} f(x;\theta) dv - \frac{d}{d\theta} \int \frac{d}{d\theta} log f(x;\theta) f(x;\theta) dv$

$$\overset{2.1}{\underset{2.4}{=}} - \int \frac{d^2}{d\theta^2} log f(x;\theta) f(x;\theta) dv = E_\theta \frac{d}{d\theta} K(\theta; X)$$

[B.Rüger (1999)](Seite 93)

Aus dem Beweis erfahren wir, dass die Varianz von Score-Funktion ist eine negative Erwartungswert von Score-Funktion über θ.

2.2.2 Fisher-Information einer Stichprobe X

Zur Definition der Fisher-Information bengen wir eine zugrunde liegende Fisher-reguläre Verteilungsannahme $\mathfrak{B} = \{f(x;\theta) : \theta \in \Theta\}$ über X. Dann ist die Fisher-Information einer Beobachtung x definiert als:

$$I(\theta;x) := -\frac{d}{d\theta}K(\theta;x) = -\frac{d^2}{d\theta^2}logf(x;\theta) \tag{2.6}$$

Wir berücksitigen den Beweis von (2.5), die Fisher-Information einer Beobachtung x $I(\theta;x)$ ist nichts mehr als die negative Differentiation der Score-Funktion. $I(\theta;x)$ bekommen wir nach der Beobachtung der Stichprobe. Die Erwartungswert der Zufallsgröße $I(\theta;X)$ nennen wir die erwartete Fisher-Information $I_X(\theta)$ über θ oder einfach die Information von der Stichprobe X. Und die Fisher-Information $I_X(\theta)$ soll vor der Beobachtung erwartet werden (a priori Standpunkt). Eine zusammenstellung von (2.4), (2,5) und (2,6) folgt:

$$I_X(\theta) = E_\theta I(\theta;X) = Var_\theta K(\theta;X) = E_\theta K^2(\theta;X) \tag{2.7}$$

[B.Rüger (1999)](Seite 93-94)

Wie oben erwähnt, die Score-Funktion $K(\theta;X)$ kann als ein Maß angesehen werden, wie gut man aufgrund einer Beobachtung x die verschiedenen Dichten lokal trennen kann. In diesem Fall, ist $I_X(\theta)$ als der Erwartungswert von $K^2(\theta;X)$ auch ein Maß dafür, wie gut man aufgrund eine beobachtete Stichprobe die Dichten lokal trennen kann. Wir wissen schon die Funktion der Fisher-Information, jedoch der Parameter θ als wahrer Wert, der wir bewerten sollen, unbekannt ist. Wir nehmen an, daß die Dichtfunktion $f(x;\theta)$ ein eindeutig bestimmtes Maximalstelle $\hat{\theta}$ für jede Beobachtung x besitzt. Durch der Logarithmierung der nomierte Dichtfunktion $f(x;\theta)$ erhalten wir eine Gleichung:

$$k(\theta;x) = -\frac{1}{2}(\theta-\hat{\theta})^2 I(\hat{\theta};x)$$

$k(\theta;x)$ hier sitmmt eine Parabel von θ überein, die ihren Scheitel an der Stelle $\hat{\theta}$ und den Krümmungsradius $1/I(\hat{\theta};x)$ hat. Für die Dichtfunktion bedeutet es: Je größer die Information $I(\hat{\theta};x)$ ist, desto stärker ist die Parabel gekrümmt und desto schärfer und genauer ist die Maximalstelle $\hat{\theta}$.

Durch Rücktransfornation auf $f(x;\theta)$ stellt eine approximative Verteilung dar:

$$f(x;\theta) \sim N(\hat{\theta}; 1/I(\hat{\theta};x))$$

Die $\hat{\theta}$ ist der Erwartungswert der approximativen Verteilung und die Inverse der Fisher-Information $I(\hat{\theta};x)$ ist die Varianz. In diesem Sinne erfassen wir, daß je größer die Fisher-Information ist ,desto kleiner ist die Varianz der Verteilung und desto genauer ist die Schätzung $\hat{\theta}$ aus der Beobachtung x für das unbekannte θ.

2.2.3 Fisher-Information einer Statistik T

Nun gehen wir auf die Fisher-Information $I_T(\theta)$ einer Statistik $T = T(X)$ ein. Stellen wir eine Verteilungsfamile $\mathfrak{B}^T = \{f_T(t;\theta)\}$ mit den Verteilungen von T dar, die durch das Bildmaß v^T dominiert werden. Aus der Fisher-Regularitätsbedingung der Verteilungsfamilie $\mathfrak{B}$ kann man sagen, dass $\mathfrak{B}^T$ auch Fisher-regulär ist. Mit der Dichtfunktion der Verteilung von T, folgt die Score-Funktion von $t = T(x)$:

$$K_T(\theta;t) = \frac{d}{d\theta} log f_T(t;\theta) \tag{2.8}$$

Deshalb ist $K_T(\theta;T)$ die Score-Funktion von T über θ und vgl. (2.7):

$$I_T(\theta) = I_{T(x)}(\theta) = Var_\theta K_T(\theta;T) = E_\theta K_T^2(\theta;T) \tag{2.9}$$

[B.Rüger (1999)](Seite 96)
In diesem Abschnitt statt x setzen wir $T(X)$ ein, erhalten wir eine Abschätzung: $I_T(\theta) \leq I_X(\theta)$ für alle θ. Diese Abschätzung ist Teil von dem Hauptsatz über die Fisher-Information. Um disen zu beweisen benötigen wir 4 besondere Aussagen.

Die erste Aussage betrifft die Vertauschbarkeitseigenschaften der Fisher-Regularität der Verteilungsfamilie $\mathfrak{B}$. Aber die Vertauschbarkeitsregeln gelten nicht nur ür die INtegrage über ganz $\mathfrak{X}$, sonder auch für die Integrale über irgendeine meßbare Teilmenge $B \subset \mathfrak{X}$. D.h. diese Vertauschbatkeitsregeln gelten auch für die Dichtfunktion $f_T(t;\theta)$ der Verteilungsfamile $\mathfrak{B}^T$.

Die zweite Aussage betrifft die Score-Funktion von T und sie steht mit der Score-Funktion der Stichprobe X in folgende Beziehung dar:

$$K_T(\theta;t) = E_\theta[K(\theta;X)|T = t] \tag{2.10}$$

Die dritte Aussage mit der Hilfe des Satzes vom iterierten Erwartungswert folgt die Kovarianz der Score-Funktion $K(\theta;X)$ und $K_T(\theta;T)$ in die folgende Beziehung:

$$E_\theta[K(\theta;X)K_T(\theta;T(X))] = E_\theta[K_T^2(\theta;T(X))]$$

$$cov_\theta[K(\theta;X), K_T(\theta;T(X))] = I_{T(X)}(\theta) \tag{2.11}$$

Die vierte Aussage ist benötigen wir für Teil des Hauptsatzes und sie lautet:

$$I_X(\theta) - I_{T(X)}(\theta) = Var_\theta[K(\theta;X) - K_T(\theta;T(X))] \tag{2.12}$$

Für den Fall $I_{T(X)}(\theta) = I_x(\theta)$ für alle θ betrachten wir eine Ausnahme einer v-Nullmenge N_θ von x. Und wenn die Ausnahmemenge N_θ unabhängig von θ ist, gilt die folgende Beziehung:

$$K(\theta;x) = K_T(\theta;T(x)) \text{ für alle } x \in \mathfrak{X}\backslash N \tag{2.13}$$

mit $\nu(N) = 0$ und $K(\theta;X) = K_T(\theta;T(X))$. [B.Rüger (1999)](Seite 97-98)

2.2.4 Hauptsatz der Fisher-Information

In diesem Abschnitt bezieht sich wieder auf eine Fisher-reguläre Verteilungsannahme $\mathfrak{B}$ über die Stichprobe X mit einer Score-Funktion $K(\theta;X)$ und einer endlichen Fisher-Information. Dann wird die Fisher-Information $I_{T(X)}(\theta)$ durch den nachfolgenden Aussagen überprüft:

a) $I_T(\theta) \leq I_X(\theta)$ für alle θ.
Beweis: Von (2.15) wird den Hauptsatz a) folglich beweist.

b) Ist T suffizient für $\mathfrak{B}$, so gilt $I_{T(X)}(\theta) = I_X(\theta)$ für alle θ.

Die Aussage b) ist unter Bezugnahme auf die Charakterisierung der Suffizienz. Wir erklären hier kurz über den Begriff Suffizienz. Suffizienz heißt: für Auswetung der Beobachtung x einer Stichprobe X, x reduziert durch einer Stichprobefunktion $T(x)$, wenn

$T(x)$ keine Verlust an Information hat. D.h. mit einer suffizient Statistik T soll die Stichprobe X ohne Informationsverlust reduzieren. (vgl.[B.Rüger (1999)] Seite 38)

Beweis: Mit Hilfe des Satzes von Nayman-Kriterium, wenn T suffizient ist, haben die Dichtfunktionen $f(x;\theta)$ und $f_T(T(x);\theta)$ eine Beziehung in:
$f(x;\theta) = h(x)f_T(T(x);\theta)$ mit $h(x) = exp\{\tilde{h}(x)\}$
Wenn wir die Score-Funktion von jeweiligen Dichtfunktionen rechnen, dann bekommen wir eine Gleichung:$K(\theta;x) = K_T(\theta;T(x))$. Mit den oben erwähnte Ausnahmemenge N_θ ist damit:
$I_{T(X)}(\theta) = Var_\theta K(\theta;X) = Var_\theta K_T(\theta;T(X)) = I_X(\theta)$ für alle θ

c) Gilt $I_{T(X)}(\theta) = I_X(\theta)$ für alle θ und zusäthlich die Annahme (2.16), so ist T suffizient für $\mathfrak{B}$.
Beweis: Aus b) wissen wir, wenn T suffizient ist, dann ist $I_{T(X)}(\theta) = I_X(\theta)$. Andersherum ist $I_{T(X)}(\theta) = I_X(\theta)$, können wir auch schätzen, dass T suffizient ist. "Eine Statistik T ist fenau dann suffizient, wenn ihre Score-Funktion mit derjanigen von X übereinstimmet.

Beispiel: Information in Bernoulli-Experiment.

Das Bernoulli-Experiment besitzt eine Verteilung der Stichprobe X: $f(x;p) = p^z(1-p)^{n-z}$ mit $z = \sum x_i$. Die Score-Funktion der Beobachtung x lautet:

$$K(p;x) = \frac{z-np}{p(1-p)}$$

Die Fisher-Information gilt:

$$I_X(p) = n/p(1-p)$$

Wir nehmen eine suffizient Statistik Z und die Verteilung der Statistik Z lautet: $f_Z(z;p) = \binom{n}{z}p^z(1-p)^{n-z}$. Nach dem Hauptsatz gilt: $I_Z(p) = I_X(p)$ für alle p gilt:

$$K_Z(p;z) = \frac{z-np}{p(1-p)} = K(p;x)$$

Die Score-Funktion von Z mit derjenigen von X übereinstimmt.
"([B.Rüger (1999)],Seite 99-100)

2.2.5 Fisher-Information mit mehrere Statistiken

In diesem Abschnitt Beweisen wir die Fisher-Information über der Stichprobe X und mehreren Statistik $T_1(X)$, $T_2(X)$ und $V(X)$:

a) Ist V ancillary für $\mathfrak{B}$, so ist $I_V(\theta) = 0$ für all θ.
Beweis: Eine ancillary Statistik V für $\mathfrak{B}$ heißt, die Verteilung von V nicht bon θ abhängt. Die Statistik V allein enthält keine Information über θ. Sodass die Dichtfunktion $f_v(v;\theta)$ unabhängig von θ ist, gilt die Score-Funktion $K(\theta;v) \equiv 0$.

b)(Additivität der Information) Sind T_1 und T_2 unab"hangig, so gilt:

$$I_(T_1, T_2)(\theta) = I_{T_1}(\theta) + I_{T_2}(\theta) \tag{2.14}$$

Beweis: In diesem Fall sollen T_1 und T_2 unabhängig sein. Weil $log(a \cdot b) = log(a) + log(b)$, führen wir die Addition der beiden unabhängigen Score-Funktionen ein:

$$K_(T_1, T_2)(\theta; (t_1, t_2)) = K_{T_1}(\theta; t_1) + K_{T_1}(\theta; t_1)$$

Damit folgt die Information:

$$\begin{aligned} I_{(T_1,T_2)}(\theta) &= Var_\theta K_{(T_1,T_2)}(\theta; (t_1, t_2)) = Var_\theta K_{T_1}(\theta; t_1) + Var_\theta K_{T_1}(\theta; t_1) \\ &= I_{T1}(\theta) + I_{T2}(\theta) \end{aligned}$$

c) Ist $X = (X_1, \ldots, X_2)$ eine i.i.d Stichprobe, so gilt für alle θ:

$$I_X(\theta) = nI_{X_1}(\theta) \tag{2.15}$$

Beweis: folgt aus b). [B.Rüger (1999)](Seite 101)

2.2.6 Fisher-Information mit speziellem Parameter

In diesem Abschnitt betrachten wir, dass die Fisher-Information auch von der speziellen Parameter γ der Fisher-reguläre Verteilungsfamile $\mathfrak{B}$ abhängt. Wir nehmen der Parameter $\gamma = \gamma(\theta)$ als eine Transformation an Stelle von θ an. In diesem Fall sind Parametern θ und $\gamma(\theta)$ äquivalent, obwohl die Information $\tilde{I}_X(\gamma) = \tilde{I}_X[\gamma(\theta)] \neq I_X(\theta)$. Die Beziehungen beiden Information soll mit einer nirgends verschwindenden Ableitung $\gamma'(\theta)$ folgen:

$$\tilde{I}_X[\gamma(\theta)] \cdot [\gamma'(\theta)]^2 = Var_\gamma(\theta)[\tfrac{d}{d\gamma(\theta)} log f(X;\gamma(\theta)) \cdot \gamma'(\theta)]$$

$$= Var_\theta[\tfrac{d}{d\theta} log f(X;\theta)] = I_X(\theta)$$

[B.Rüger (1999)](Seite 101-102)

3 Informationsmatrix

Im letzten Kapitel haben wir die Information einer Fisher-reguläre Verteilungsannahme $\mathfrak{B}$ über eine unbekannte Parameter θ erstellt. Im diesem Kapitel betrefft einer mehrdimensionaler Parametervektor $\theta = (\theta_1, \ldots, \theta_r), r \geq 2$. Zu den Beweise der Aussagen in diesem Kapital bentigen wir teilweise von oben erwähnte Schritte.

3.1 Informationsmatrix

3.1.1 Fishersche Informationsmatrix einer Stichprobe

Wie obige Information einer Stichprobe X über Paremeter θ benötige wir in diesem Kapotel auch einer voraussetzende Fisher-reguläre Verteilungsfamilie. Die Regularitätseigenschaften sollen sich in beiden Kapitale sehr nähre kommen. Sieht den Abschnitt 2.1. Nun gilt die Vertauschbarkeitsregeln auch für die Dichtfunktion $f(x;\theta)$ mit der betreffenden Parametervektor $\theta = (\theta_1, \ldots, \theta_r), r \geq 2$. D.h. für $i = 1, \ldots, r$ und $j = 1, \ldots, r$ gilt die Vertauschbarkeitsregeln in folgenden Beziehungen:

$$\int \frac{\partial}{\partial \theta_i} f(x;\theta) dv = \frac{\partial}{\partial \theta_i} \int f(x;\theta) dv = 0 \tag{2.16}$$

$$\int \frac{\partial^2}{\partial \theta_i \partial \theta_j} f(x;\theta) dv = \frac{\partial}{\partial \theta_j} \int \frac{\partial}{\partial \theta_i} f(x;\theta) dv = 0 \tag{2.17}$$

Damit ist die r-dimensionale Score-Vektor $K(\theta;x)$ der Stichprobe X:

$$K(\theta;x) = (K_1(\theta;X), \ldots, K_r(\theta;X)) \quad mit \tag{2.18}$$

$$K_i(\theta;x) = \frac{\partial}{\partial \theta_i} log f(x;\theta), i = 1, \ldots, r$$

Durch Kovarianz der Score-Funktion bekommen bekommen:

$$cov_\theta[K_i(\theta;X),K_j(\theta;X)] = -E_\theta\frac{\partial}{\partial\theta_i}K_j(\theta;X) = -E_\theta\frac{\partial^2}{\partial\theta_i\partial\theta_j}logf(x;\theta) \tag{2.19}$$

Im Bezugnahme auf den ein-dimensionale Fall von (1.1) und(1.2) wird den nächste Schritt analog beweist. Wir nennen die Zufallsgröße $I(\theta;x)$ der symmertischen $(r \times r)$-Matrix die Fishersche-Informationsmatrix über der Beobachtung x über den Parameter θ.

$$I(\theta;x) = (-\frac{\partial^2}{\partial\theta_i\partial\theta_j}logf(x;\theta))_{i=1,\dots,r,j=1,\dots,r} \tag{2.20}$$

Der Erwartungswert dieser Matix nennen wir die erwartete Fishersche-Informationsmatrix $I_X(\theta)$. Die bezieht sich auf die Situation vor der Beobachtung. Es folgt:

$$I_X(\theta) = E_\theta I(\theta;X) = Cov_\theta K(\theta;X) \tag{2.21}$$

[B.Rüger (1999)](Seite 104)

3.1.2 Fishersche Informationsmatrix einer Statistik

Für weitere Beweise einer Statistik $T = T(X)$ sollen analog wie oben sein.

Die r-dimensionale Score-Vektor der T mit $t = T(x)$ gilt:

$$K_T(\theta;t) = (K_{T;1}(\theta;t),\dots,K_{T;r}(\theta;t)) \quad mit \tag{2.22}$$

$$K_{T;i}(\theta;t) = \frac{\partial}{\partial\theta_i}logf_T(t;\theta), i = 1,\dots,r$$

Die Kovarianz der Score-Vektor und die Fisherscher Informationsmatrix von T stellt die Beziehung dar:

$$I_T(\theta) = I_{T(X)}(\theta) = Cov_\theta K_T(\theta;T) \tag{2.23}$$

Nach der Vergleichung der Informationsmatrizen $I_X(\theta)$ und $I_{T(X)}(\theta)$ folgt:

$$I_T(\theta) - I_{T(X)}(\theta) = Cov_\theta[K(\theta;X) - K_T(\theta;T(X))] \tag{2.24}$$

Mit eine Ausnahmemenge vgl. analog Beweis aus Abschnitt 2.23 von (2.15) und (2.16) setllen wir die Annahme dar:

$$K(\theta; x) = K_T(\theta; T(x)) \quad \textit{für alle } x \in \mathfrak{X} \backslash N$$

[B.Rüger (1999)](Seite 105)

Der Hauptsatz der Fisher-Information bleibt auch für Informationsmatrix gültig. Hier werden die Aussagen und Beweisen verzichten. Sieht Abschnitt 2.2.4.

Beispiel: Informationsmatrix im Gauß-Experiment.

Wir betrachten das Gauß-Experiment mit dem zweidimensionalen Parameter $\theta = (\mu, \sigma^2)$, wobei μ und σ^2 unbekannt sind, sodaß die Informationsmatrix eine (2×2)-Matrix ist. Die Dichte der Gauß-Experiment lautet: $f(x;\theta) = (\frac{1}{\sqrt{2\pi\sigma^2}})^n exp\left\{-\frac{1}{2\sigma^2}\sum(x_i-\mu)^2\right\}$.
Die Score-Vektor gilt: $K(\theta; X) = (K_1(\theta; X), K_2(\theta; X))$ mit der Komponenten:

$K_1(\theta; X) = \frac{\partial}{\partial\mu} log f(X;\theta) = \frac{n}{\sigma^2}(\bar{X} - \mu)$ und
$K_2(\theta; X) = \frac{\partial}{\partial\sigma^2} log f(X;\theta) = \frac{n}{2\sigma^4}(S_\mu^2 - \sigma^2)$

Die Varianzen sind jeweils: $Var_\theta K_1(\theta; X) = \frac{n}{\sigma^2}$ und $Var_\theta K_2(\theta; X) = \frac{n}{2\sigma^4}$

Da $\bar{X}$ und S^2 unabhängig sind, ist $cov_\theta(K_1(\theta; X), K_2(\theta; X)) = 0$.

Die Fishersche Informationsmatrix folgt: $I_X(\mu, \sigma^2) = \begin{pmatrix} n/\sigma^2 & 0 \\ 0 & n/2\sigma^4 \end{pmatrix}$

[B.Rüger (1999)](Seite 107)

Literaturverzeichnis

[B.Rüger (1999)] Dr. Bernhard Rüger, *Test- und Schätztheorie Band I: Grundlagen.* R.Oldenbourg Verlag München Wien, 1999.