02.12.10

Fachbereich Mathematik und Informatik
Institut für Mathematische Statistik
Seminar: Große Abweichungen
Dozent: Prof. Dr. Matthias Löwe
Wintersemester 2010/2011

Ausarbeitung zum Seminarvortrag

Der Satz von Gärtner-Ellis

vom 4. Dezember 2010

Sabrina Schramm
sabrina.schramm@uni-muenster.de
Matrikelnummer: 349701

Inhaltsverzeichnis

1 Einleitung

Die Theorie der Großen Abweichungen ist ein Teilgebiet der Wahrscheinlichkeitstheorie, das sich mit dem Studium sehr seltener Ereignisse beschäftigt. In allen einführenden Texten über die Wahrscheinlichkeitstheorie werden asymptotische Aussagen wie das Schwache und das Starke Gesetz der Großen Zahlen oder der Zentrale Grenzwertsatz behandelt. Diese beschreiben das normale Verhalten einer Partialsummenfolge von unabhängig, identisch verteilten Zufallsgrößen. Die Theorie der Großen Abweichungen möchte im Gegenzug das Abweichungsverhalten analysieren. Dabei spielt unter Anderem die Betrachtung der Wahrscheinlichkeit der Ereignisse $\{\frac{1}{n}S_n > x\}$ beziehungsweise $\{\frac{1}{n}S_n < -x\}$ eine zentrale Rolle. Nun stellt sich vielleicht die Frage, warum die Theorie der Großen Abweichungen eine so große Rolle spielt. Die Liste der Anwendungen dieser Theorie zumindest ist lang und wächst ständig. Als Beispiel seien an dieser Stelle nur die Statistik und die Informationstheorie genannt. Aber bei Infragestellen dürfte allein ein Gedanke an die geringe Wahrscheinlichkeit, im Lottospiel Sechs Richtige zu tippen, zunächst einmal etwaige Zweifel beiseite räumen.
Das Thema meines Seminarvortrages - Der Satz von Gärtner-Ellis - stellt eine weit reichende Verallgemeinerung des Satzes von Cramér dar, der sich innerhalb der Theorie lediglich auf den Fall von unabhängig, identisch verteilten Zufallsvariablen beschränkt. In meinem Vortrag werde ich zunächst kurz darauf eingehen, was genau unter einem Prinzip Großer Abweichungen verstanden wird. Davon ausgehend werde ich nach einer Einführung grundlegender Definitionen und Voraussetzungen den Satz von Gärtner-Ellis formulieren und beweisen.

2 Das Prinzip Großer Abweichungen

Seien X_1, X_2, ..., X_n unabhängig, identisch verteilte Zufallsgrößen, deren Varianz σ^2 existiert und $\mathrm{E}(X_1) = 0$ gelte. Wir betrachten die Partialsummenfolge $(S_n)_{n\in\mathbb{N}}$ und möchten ihr Verhalten für große n diskutieren. Dabei liegt unsere Betrachtungsweise auf der Analyse des Abweichungsverhaltens, also der Wahrscheinlichkeiten sehr seltener Ereignisse.
Zunächst wollen wir definieren, was wir unter einem Prinzip Großer Abweichungen verstehen:

Es sei (E,d) ein metrischer Raum und B_E die Borel-σ-Algebra auf E, d.h. die von den offenen Mengen erzeugte σ-Algebra. Ferner seien $(\mu_n)_{n\in\mathbb{N}}$ eine Folge von Wahrscheinlichkeitsmaßen auf (E,B_E) und $(\gamma_n)_{n\in\mathbb{N}}$ eine Folge positiver Zahlen mit $\gamma_n \to \infty$.

Definition 2.1. (Prinzip Großer Abweichungen): Wir sagen, die Folge $(\mu_n)_{n\in\mathbb{N}}$ genügt einem Prinzip Großer Abweichungen mit Ratenfunktion I und Skala γ_n, falls eine Funktion I: E $\to$ [0,∞] existiert mit:

1. 0 $\leq$ I(x) $\leq$ ∞ für alle x $\in$ E
2. I ist von unten halbstetig, d.h. für jede Folge $(x_n)_{n\in\mathbb{N}}$ in E und für jedes x $\in$ E gilt: ($x_n$$\to$x) $\to$ ($\liminf_{n\to\infty}$ I(x_n) $\geq$ I(x)).
3. Für jedes I < ∞ ist die Menge { x $\in$ E : I(x) $\leq$ s }, s $\geq$ 0 kompakt in E.
4. Für jede abgeschlossene Menge F $\subset$ E gilt:
   $\limsup_{n\to\infty} \frac{1}{\gamma_n}$ log μ_n (F) $\leq$ - $\inf_{x\in F}$ I(x)
5. Für jede offene Menge G $\subset$ E gilt:
   $\liminf_{n\to\infty} \frac{1}{\gamma_n}$ log μ_n (G) $\geq$ - $\inf_{x\in G}$ I(x)

Bemerkung. Die Bedingung 4. nennen wir auch obere Schranke für abgeschlossene Mengen, Bedingung 5. analog untere Schranke für offene Mengen.

3 Grundlegende Definitionen und Voraussetzungen

Definition 3.1. Es sei (Z_n) eine Folge von reellwertigen, borelmessbaren Zufallsvektoren auf dem Wahrscheinlichkeitsraum (Ω,B($\mathbb{R}$),P). Dann ist die Momentenerzeugende Funktion definiert als

$$\phi_n(t) := \mathbb{E}(\exp(< t, Z_n >)), t \in \mathbb{R}^d, n \in \mathbb{N}$$

Ferner definieren wir die logarithmische Momentenerzeugende Funktion als

$$\log \phi_n(t) := \log(\mathbb{E}(\exp(< t, Z_n >))), t \in \mathbb{R}^d, n \in \mathbb{N}$$

Die Fenchel-Legendre-Transformierte ist definiert als

$$\Lambda^*(x) := \sup_{\lambda\in\mathbb{R}^d}\{< x, \lambda > -\Lambda(\lambda)\} \text{ mit } \Lambda(\lambda) := \lim_{n\to\infty} \frac{1}{n} \log \phi_n(n\lambda)$$

Definition 3.2. (exponierter Punkt) Ein $y \in \mathbb{R}^d$ heißt exponierter Punkt von Λ^*, wenn ein $\lambda \in \mathbb{R}^d$ existiert, sodass für alle x $\neq$ y gilt:

$$< \lambda, y > -\Lambda^*(y) >< \lambda, x > -\Lambda^*(x)$$

Ein solches λ heißt exponierende Ebene zu y.

These. 1. Für jedes $\lambda \in \mathbb{R}^d$ existiere $\Lambda(\lambda) := \lim_{n\to\infty} \frac{1}{n}$ log ϕ_n(nλ) $\in$ [-∞,∞]

2. Es gelte 0 $\in$ int(D_Λ), wobei $D_\Lambda := \{\lambda \in \mathbb{R}^d : \Lambda\ (\lambda) < \infty\}$

Lemma 3.3. *Λ hat die Eigenschaften:*

1. *Λ ist konvex und Λ > - ∞*

2. *Λ^* ist eine Ratenfunktion und konvex.*

Definition 3.4. (Relatives Inneres) Für jede nicht-leere Menge C $\subseteq \mathbb{R}^d$ ist das relative Innere von C definiert als die Menge

$$ri(C) := \{y \in C : x \in C \to y - \epsilon(x-y) \in C \text{ für ein } \epsilon > 0\}$$

Bemerkung.

- Das relative Innere ist sozusagen das Innere aus der Menge selbst heraus betrachtet.
- Es gilt: int(A) $\subset$ ri(A).
- Wenn A = {x}, dann ist int(A) = $\emptyset$, während ri(A) = A.

Definition 3.5. (Exponentielle Straffheit) Eine Familie $(\mu_n)_{n\in\mathbb{N}}$ von Wahrscheinlichkeitsmaßen auf (E,B_E) heißt exponentiell straff auf der Skala (γ_n), falls es zu jedem ϵ > 0 eine kompakte Menge K gibt, sodass

$$\limsup_{n\to\infty} \frac{1}{\gamma_n} \log \mu_n(K^c) \leq -\epsilon$$

Lemma 3.6. *Für alle Folgen $(a_n)_{n\in\mathbb{N}}$ und $(b_n)_{n\in\mathbb{N}}$ in (0,∞) gilt:*

$$\lim_{n\to\infty} \frac{1}{n} \log(a_n + b_n) = \lim_{n\to\infty} \frac{1}{n} \log(\max\{a_n, b_n\})$$

Definition 3.7. Definiere Λ^*(S) := $\inf_{x\in S}$ Λ^*(x) mit S $\subset \mathbb{R}^d$.

4 Das Gärtner-Ellis Theorem

Satz 4.1. *(Gärtner-Ellis): Wir setzen die obige These als gültig voraus. Dann gilt:*

1. $\limsup_{n\to\infty} \frac{1}{n}$ *log μ_n(F) $\leq$ - $\inf_{x\in F}$ Λ^*(x) für alle F $\in \mathbb{R}^d$ abgeschlossen*

2. $\liminf_{n\to\infty} \frac{1}{n}$ *log μ_n(G) $\geq$ - $\inf_{x\in G\cap F}$ Λ^*(x) für alle G $\in \mathbb{R}^d$ offen, wobei F die Menge der exponierten Punkte von Λ^* bezeichnet, deren exponierende Hyperebene zu D^0_Λ gehört.*

3. *Gilt zusätzlich für Λ:*

(a) Λ ist auf $\mathbb{R}^d$ von unten halbstetig.

(b) Λ ist differenzierbar im Inneren von D_Λ.

(c) Gilt entweder $D_\Lambda = \mathbb{R}^d$ oder Λ ist steil, d.h. $\lim_{n\to\infty} |\Lambda(\lambda_n)| = \infty$ für jede Folge (λ_n) im Inneren von D_Λ, die gegen den Rand von D_Λ konvergiert.

Dann können wir G ∩ F durch G ersetzen und μ_n genügt einem Prinzip Großer Abweichungen auf $\mathbb{R}^d$ mit Rate n und Ratenfunktion Λ^*.

Beweis. Der Beweis des Satzes gliedert sich in drei Teile. Zunächst wird die obere Schranke, danach die untere Schranke und schließlich der Übergang von G ∩ F zu G gezeigt.

Teil 1: obere Schranke

Wir beweisen die Behauptung zunächst für kompakte Mengen. Sei dazu $\epsilon > 0$ beliebig und Λ^*_ϵ := min $\{\Lambda^*$(y) - ϵ, $\frac{1}{\epsilon}\}$ für y $\in \mathbb{R}^d$. Es handelt sich also um eine durch $\frac{1}{\epsilon}$ nach oben beschränkte, sich an Λ^* annähernde Minorante von Λ^*.
Da Λ^*_ϵ(y) $\leq^{Minorante}$ Λ^*(y) $=^{Def.}$ $sup_{\lambda\in\mathbb{R}^d}$ (<y,λ> - $\Lambda(\lambda)$), existiert für jedes y ein λ_y $\in \mathbb{R}^d$, sodass

$$< y, \lambda_y > -\Lambda(\lambda_y) \geq \Lambda^*_\epsilon(y) \iff \Lambda^*_\epsilon(y) \leq < y, \lambda_y > -\Lambda(\lambda_y) \quad (1)$$

Ferner finden wir für jedes y $\in \mathbb{R}^d$ eine Umgebung A_y von y, sodass

$$\inf_{x\in A_y} < x, \lambda_y > \geq < y, \lambda_y > -\epsilon \iff \inf_{x\in A_y} < x - y, \lambda_y > \geq -\epsilon \text{ gilt.} \quad (2)$$

Wir leiten nun die exponentielle Tschebyscheff- Ungleichung für eine Zufallsvariable Y her:

$$\begin{aligned} P(Y \geq \delta) &= P(\exp(nY) \geq \exp(n\delta)) \\ &= P(\exp(nY - n\delta) \geq 1) \\ &= P(\exp(n(Y - \delta)) \geq 1) \\ &= P(|\exp(n(Y - \delta)) - 0| \geq 1) \\ &\leq \mathbb{E}(\exp(n(Y - \delta))) \end{aligned}$$

und wenden diese an. Dann erhalten wir mit (2):

$$\begin{aligned}
\mu_n(A_y) &= P(Z_n \in A_y) \\
&\leq^{(2)} P(< Z_n - y, t_y > \geq -\epsilon) \\
&= P(\exp(n < Z_n - y, t_y >) \geq \exp(-n\epsilon)) \\
&= P(\exp(n < Z_n - y, t_y > +n\epsilon) \geq 1) \\
&\leq^{Tscheby} \mathbb{E}(\exp(n < Z_n - y, t_y > +n\epsilon)) \\
&= \exp(n\epsilon)\mathbb{E}(\exp(n < Z_n - y, t_y >)) \\
&= \exp(n\epsilon)\exp(-n < y, t_y >)\mathbb{E}(\exp(n < t_y, Z_n >)) \\
&= \exp(n\epsilon)\exp(-n < y, t_y >)\mathbb{E}(\exp(< nt_y, Z_n >)) \\
&=^{Def.} \exp(\epsilon n)\exp(-n < y, t_y >)\phi_n(nt_y)
\end{aligned}$$

Sei K $\subset \mathbb{R}^d$ eine kompakte Menge. Dann hat die Überdeckung $\bigcup_{y \in K} A_y$ von K eine endliche Teilüberdeckung $\bigcup_{i=1,\ldots,N} A_{y_i}$ und wir erhalten mit der obigen Abschätzung:

$$\begin{aligned}
\frac{1}{n}\log\mu_n(K) &\leq \frac{1}{n}\log[N \max_{i=1,\ldots,N}\mu_n(A_i)] \\
&= \frac{1}{n}\log[N \max_{i=1,\ldots,N}(\exp(\epsilon n)\exp(-n < y, t_{y_i} >)\phi_n(nt_{y_i}))] \\
&= \frac{1}{n}\log N + \frac{1}{n}\log\exp(\epsilon n) + \max_{i=1,\ldots,N}(\frac{1}{n}(-n < y, t_{y_i} >) + \frac{1}{n}\log\phi_n(nt_{y_i})) \\
&= \frac{1}{n}\log N + \epsilon + \max_{i=1,\ldots,N}(- < y, t_{y_i} > + \frac{1}{n}\log\phi_n(nt_{y_i})) \\
&= \frac{1}{n}\log N + \epsilon - \min_{i=1,\ldots,N}(< y, t_{y_i} > - \frac{1}{n}\log\phi_n(nt_{y_i}))
\end{aligned}$$

N ist konstant und für $n \to \infty$ erhalten wir:

$$\begin{aligned}
\limsup_{n\to\infty}\frac{1}{n}\log\mu_n(K) &\leq \epsilon - \min_{i=1,\ldots,N}(< y, t_{y_i} > - \limsup_{n\to\infty}\frac{1}{n}\log\phi_n(nt_{y_i})) \\
&= \epsilon - \min_{i=1,\ldots,N}(< y, t_{y_i} > - \Lambda(t_{y_i})) \\
&\leq^{(1)} \epsilon - \min_{i=1,\ldots,N}\Lambda^*_\epsilon(y_i) \\
&\leq \epsilon - \Lambda^*_\epsilon(K)
\end{aligned}$$

Für $\epsilon \to 0$ folgt nun wegen Λ^*_ϵ(y) $\to \min$ $\{\Lambda^*$(y) - 0, $\infty\}$ die gewünschte Abschätzung für kompakte Teilmengen:

$$\limsup_{n\to\infty}\frac{1}{n}\log\mu_n(K) \leq -\Lambda^*(K) \tag{3}$$

Die Erweiterung kompakter Mengen zu abgeschlossenen ergibt sich nun durch den Beweis der exponentiellen Straffheit von $(\mu_n)_{n\in\mathbb{N}}$. Ist C eine abgeschlossene Teilmenge des $\mathbb{R}^d$, dann ist C $\cap$ $[-N, N]^d$ kompakt für alle N $>$ 0.

- Vollständigkeit wird durch den $\mathbb{R}^d$ gewährleistet.
- Die Beschränktheit ergibt sich durch den Schnitt.

Wir müssen noch zeigen, dass die Ungleichung für den Grenzwert N $\to \infty$ gilt, da wir wissen, dass $\lim_{N\to\infty}$ (C $\cap$ $[-N,N]^d$) = C.
Sei dazu N' := $[-N,N]^d$. Dann gilt:

$$\begin{aligned}
\limsup_{n\to\infty}\frac{1}{n}\log\mu_n(C) &\leq^1 \limsup_{n\to\infty}\frac{1}{n}\log\mu_n[(C\cap N')\cup(\mathbb{R}^d-N')] \\
&= \limsup_{n\to\infty}\frac{1}{n}\log(\mu_n[(C\cap N')]+\mu_n[(\mathbb{R}^d-N')]) \\
&=^{\text{Lemma 3.5.}} \limsup_{n\to\infty}\frac{1}{n}\log\max\{\mu_n[(C\cap N')],\mu_n[(\mathbb{R}^d-N')]\} \\
&=^{(3)} \max\{-\Lambda^*(C\cap N'),\limsup_{n\to\infty}\frac{1}{n}\log\mu_n(\mathbb{R}^d-[-N,N]^d)\}
\end{aligned}$$

für alle N > 0
1 (C $\cap$ N'), ($\mathbb{R}^d$ - N') sind disjunkt und enthalten zusammen ganz C

Es fehlt nun noch zu zeigen, dass $\lim_{N\to\infty}\limsup_{n\to\infty}\frac{1}{n}$ log μ_n ($\mathbb{R}^d$ - $[-N,N]^d$) = -∞ gilt, denn damit ist dann Teil 1) bewiesen.
Seien dazu u_j die j-ten Einheitsvektoren im $\mathbb{R}^d$ für j = 1, ..., d. Da wir die These als gültig voraussetzen, finden wir einen Vektor nahe bei Null, sodass gilt:

$$\exists\delta>0:\forall i\Lambda(\pm\delta u_i)<\infty.$$

Wir wenden nun die exponentielle Tschebyscheff-Ungleichung an und es folgt für alle i:

$$\begin{aligned}
P(Z_n^{(i)}\leq -N) &= P(\exp(n\delta(-Z_n^{(i)}))\geq\exp(n\delta N)) \\
&= P(\exp(-n\delta Z_n^{(i)}-n\delta N)\geq 1) \\
&\leq^{Tscheby} \mathbb{E}(\exp(-n\delta Z_n^{(i)}-n\delta N)) \\
&= \exp(-n\delta N)\mathbb{E}(\exp(-n\delta Z_n^{(i)})) \\
&= \exp(-n\delta N)\mathbb{E}(\exp(<-n\delta,Z_n^{(i)}>)) \\
&=^{Def.} \exp(-n\delta N)\phi_n(-n\delta u_i)
\end{aligned}$$

Analog erhält man:

$$\begin{aligned} P(Z_n^{(i)} \geq N) &= P(\exp(n\delta Z_n^{(i)}) \geq \exp(n\delta N)) \\ &= P(\exp(n\delta Z_n^{(i)} - n\delta N) \geq 1) \\ &\leq^{Tscheby} \mathbb{E}(\exp(n\delta Z_n^{(i)} - n\delta N)) \\ &= \exp(-n\delta N)\mathbb{E}(\exp(n\delta Z_n^{(i)})) \\ &= \exp(-n\delta N)\phi_n(n\delta u_i) \end{aligned}$$

Dadurch können wir nun abschätzen, wie wahrscheinlich es ist, dass eine Komponente der Zufallsvariable nicht in das Intervall [-N,N] fällt. $\limsup_{n\to\infty} \frac{1}{n} \log \mu_n$ ($\mathbb{R}^d$ - $[-N,N]^d$) beschreibt hingegen die Wahrscheinlichkeit, dass die ganze Zufallsvariable außerhalb liegt. Diese ist sicherlich kleiner, als wenn $\geq$ 1 Komponenten außerhalb liegen:

$$\begin{aligned} \limsup_{n\to\infty} \frac{1}{n} \log \mu_n(\mathbb{R}^d - [-N,N]^d) &\leq \limsup_{n\to\infty} \frac{1}{n} \log P(\exists i \in \{1,...d\} : Z_n^{(i)} \notin [-N,N]) \\ &\leq^{3.5.} \max_{i=1,...d} \limsup_{n\to\infty} \frac{1}{n} \log \mu_n(\max\{\exp(-n\delta N)\phi_n(-n\delta u_i), \\ &\qquad \exp(-n\delta N)\phi_n(n\delta u_i)\}) \\ &=^{Def.} -\delta N + \max_{i=1,...d} \max\{\Lambda(-\delta u_i), \Lambda(\delta u_i)\} \end{aligned}$$

Da $\Lambda(\pm\ \delta u_i) < \infty$ vorausgesetzt wurde, folgt nun:

$$\limsup_{n\to\infty} \frac{1}{n} \log \mu_n(\mathbb{R}^d - [-N,N]^d) < -\infty$$

Das zeigt uns die exponentielle Straffheit und beendet den Beweis für die obere Schranke.

Teil 2: untere Schranke

Sei B_ϵ(y) ein offener Ball um y mit Radius ϵ. Es reicht aus, zu zeigen:

$$\lim_{\epsilon\to 0} \liminf_{n\to\infty} \frac{1}{n} \log \mu_n(B_\epsilon(y)) \geq -\Lambda^*(y)\ \forall y \in F$$

Da für jede offene Teilmenge G gilt:

$$\mu_n(G) \geq \mu_n(B_\epsilon(y))\ \forall y \in G \cap F\ \forall \epsilon \leq \epsilon_0(y)$$

mit $\epsilon_0 > 0$ so klein, dass B_{ϵ_0}(y) $\subset$ G, erhalten wir die Behauptung für $n \to \infty$, $\epsilon \to 0$ und wenn wir das Infimum über alle x $\in G \cap F$ bilden.
Sei nun ein y $\in$ F und $\eta \in$ int(D_Λ) eine exponierende Hyperebene zu y. Dann gilt:

$$\Lambda(\eta) := \lim_{n\to\infty} \frac{1}{n} \log \phi_n(n\eta) <^{Vor.} \infty$$

und es folgt $\lim_{n\to\infty} \phi_n(n\eta) < \infty$.
Also ist ϕ_n $(n\eta) < \infty$ für ein genügend großes n und wir können ein assoziiertes Wahrscheinlichkeitsmaß μ'_n definieren (analog wie im Beweis vom Satz von Cramér):

$$\frac{d\mu'_n}{d\mu_n}(x) = \frac{1}{\phi_n(n\eta)} \exp(n < x, \eta >) \tag{4}$$

mit y $\in \mathbb{R}^d$ und leiten damit eine Abschätzung von unten her:

$$\begin{aligned}
\frac{1}{n}\log \mu_n(B_\epsilon(y)) &= \frac{1}{n}\log \int_{B_\epsilon(y)} \mu_n(dx) \\
&=^{(4)} \frac{1}{n}\log \int_{B_\epsilon(y)} \phi_n(n\eta)\exp(-n < x, \eta >)\mu'_n(dx) \\
&= \frac{1}{n}\log \phi_n(n\eta) + \frac{1}{n}\log \int_{B_\epsilon(y)} \exp(-n < x, \eta >)\mu'_n(dx) \\
&\geq^1 \frac{1}{n}\log \phi_n(n\eta) + \frac{1}{n}\log[\exp(-n(< y, \eta > +\epsilon|\eta|))] \int_{B_\epsilon(y)} \mu'_n(dx) \\
&= \frac{1}{n}\log \phi_n(n\eta) - < y, \eta > -\epsilon|\eta| + \frac{1}{n}\log \mu'_n(B_\epsilon(y))
\end{aligned}$$

1 $<x,\eta> = <y,\eta> + <x - y,\eta> \leq <y,\eta> + \epsilon|\eta|$, weil x $\in B_\epsilon$(y)

Zusammenfassend erhalten wir:

$$\begin{aligned}
\lim_{\epsilon\to 0}\liminf_{n\to\infty} \frac{1}{n}\log \mu_n(B_\epsilon(y)) &\geq [\liminf_{n\to\infty} \frac{1}{n}\log \phi_n(n\eta) - < y, \eta >] + \lim_{\epsilon\to 0}\liminf_{n\to\infty} \frac{1}{n}\log \mu'_n(B_\epsilon(y)) \\
&=^{Def.} [\Lambda(\eta) - < y, \eta >] + \lim_{\epsilon\to 0}\liminf_{n\to\infty} \frac{1}{n}\log \mu'_n(B_\epsilon(y))
\end{aligned}$$

Nach Teil 1 gilt:

$$\Lambda^*(y) \geq [< y, \eta > -\Lambda(\eta)] \iff [\Lambda(\eta) - < y, \eta >] \geq -\Lambda^*(y)$$

und daraus folgt:

$$\lim_{\epsilon\to 0}\liminf_{n\to\infty} \frac{1}{n}\log \mu_n(B_\epsilon(y)) \geq -\Lambda^*(y) + \lim_{\epsilon\to 0}\liminf_{n\to\infty} \frac{1}{n}\log \mu'_n(B_\epsilon(y))$$

Es fehlt nun noch zu zeigen, dass $\lim_{\epsilon\to 0} \liminf_{n\to\infty} \frac{1}{n} \log \mu'_n(B_\epsilon(\text{y})) = 0$ gilt, denn dann ist die Behauptung gezeigt. Im Gegensatz zum unabhängig, identisch verteilten Fall benötigen wir nun ein anderes Argument:

Wir bilden die Momentenerzeugende Funktion ϕ'_n von μ'_n und erhalten:

$$\begin{aligned}
\phi'_n(n\lambda) &= \int \exp(< n\lambda, x >)\mu'_n(dx) \\
&=^{Def.} \int (\exp(< n\lambda, x >)\frac{1}{\phi_n(n\eta)} \exp(n < x, \eta >)\mu_n(dx)) \\
&= \frac{\int (\exp(n < \lambda + \eta, x >)\mu_n(dx))}{\phi_n(n\eta)} \\
&= \frac{\phi_n(n(\lambda + \eta))}{\phi_n(n\eta)}
\end{aligned}$$

Da $\Lambda'(\lambda) =^{Def.} \lim_{n\to\infty} \frac{1}{n} \log \phi'_n(\mathrm{n}\lambda) = \lim_{n\to\infty} \frac{1}{n} (\log(\frac{\phi_n(n(\lambda+\eta))}{\phi_n(n\eta)})) = \lim_{n\to\infty} \frac{1}{n}$ ($\log$ $\phi_n(\mathrm{n}(\lambda + \eta))$ - $\log \phi_n(\mathrm{n}\eta)) =^{Def.} \Lambda(\lambda + \eta)$ - $\Lambda(\eta)$, nach Konstruktion $\Lambda(\eta) < \infty$ und nach Voraussetzung $\Lambda >$ - ∞, erfüllt ϕ'_n die Bedingung 1) der These. $\Lambda'(\lambda)$ erfüllt die Bedingung 2), da $\Lambda'(0) = \Lambda(0 + \eta)$ - $\Lambda(\eta) = \Lambda(\eta)$ - $\Lambda(\eta) = 0$ und $\Lambda(\mathrm{x})$ konvex ist. Damit erhalten wir folgende Relation:

$$\begin{aligned}
\Lambda'^*(y) &:=^{\text{Teil 1}} \sup_{\lambda\in\mathbb{R}^d}[< y, \lambda > -\Lambda'(\lambda)] \\
&=^{Def.} \sup_{\lambda\in\mathbb{R}^d}[< y, \lambda > -\Lambda(\lambda + \eta) + \Lambda(\eta)] \\
&=^{\lambda':=\lambda+\eta} \sup_{\lambda'\in\mathbb{R}^d}[< y, \lambda' - \eta > -\Lambda(\lambda')] + \Lambda(\eta) \\
&= \sup_{\lambda'\in\mathbb{R}^d}[< y, \lambda' > -\Lambda(\lambda')]- < y, \eta > +\Lambda(\eta) \\
&=^{Def.} \Lambda^*(y)- < y, \eta > +\Lambda(\eta)
\end{aligned} \tag{5}$$

Lemma 3.3. sagt uns, dass Λ'^* eine Ratenfunktion ist. Also können wir Teil 1 des Satzes anwenden und erhalten:

$$\limsup_{n\to\infty} \frac{1}{n} \log \mu'_n(\mathbb{R}^d - B_\epsilon(y)) \leq - \inf_{x\in(\mathbb{R}^d - B_\epsilon(y))} \Lambda'^*(x) =^{Def.} -\Lambda'^*(\mathbb{R}^d - B_\epsilon(y))$$

für alle $\epsilon > 0$.

Λ'^* besitzt kompakte Niveaumengen und nimmt wegen der Unterhalbstetigkeit auf diesen ihre Minima an. Dann gilt mit Definition 3.6.: $\Lambda'^*(\mathbb{R}^d$ - $B_\epsilon(\mathrm{y})) = \Lambda'^*(y_0)$ für ein $y_0 \neq$ y. Ferner wissen wir, dass y $\in$ F ein exponierter Punkt ist und $\eta \in \mathrm{int}(D_\Lambda)$ eine exponierende Hyperebene für y. Wir wenden nun die Definition eines exponierten Punktes mit y_0 auf (5) an:

$$\begin{aligned}
\Lambda'^*(y_0) &=^{(5)} \Lambda^*(y_0)- < y_0, \eta > +\Lambda(\eta) \\
&\geq^1 [\Lambda^*(y_0)- < y_0, \eta >]+ < y, \eta > -\Lambda^*(y) >^2 0
\end{aligned}$$

1 da Λ^* (y) $\geq^{\text{Teil 1}}$ <y,η> - Λ (η) $\iff$ Λ (η) $\geq$ <y,η> - Λ^*(y)

2 da Λ^* (y_0) - Λ^*(y) > <y_0 - y, η> für alle $y_0 \in \mathbb{R}^d$, $y_0 \neq$ y

Damit folgt: $\limsup_{n\to\infty} \frac{1}{n} \log \mu'_n(\mathbb{R}^d$ - B_ϵ(y)) $\leq$ - $\Lambda'^*(\mathbb{R}^d$ - B_ϵ(y)) = - $\Lambda'^*(y_0)$ < 0 für alle ϵ > 0 und da für jedes n $\mu'_n(\mathbb{R}^d)$ = 1, folgt $\lim_{\epsilon\to 0} \liminf_{n\to\infty} \frac{1}{n} \log \mu'_n(B_\epsilon$(y)) = 0.

Teil 3: Übergang von G $\cap$ F zu G

Lemma. (Rockafellar) Wenn Λ die Bedingungen 3a) - c) des Satzes von Gärtner-Ellis erfüllt, dann gilt: ri(D_{Λ^*}) $\subset$ F.

Bemerkung. Nach Lemma 3.3. ist Λ^* konvex, damit ist D_{Λ^*} eine konvexe Menge. Λ^* ist eine Ratenfunktion, also konstant ungleich ∞. Damit ist $D_{\Lambda^*} \neq \emptyset$ und ri(D_{Λ^*}) $\neq \emptyset$.

Wir wollen nun zeigen, dass Λ^*(G $\cap$ F) = Λ^*(G) gilt.
Es gilt: $inf_{x\in(G\cap ri(D_{\Lambda^*}))}\Lambda^*$(x) $\geq$ $inf_{x\in G\cap F}\Lambda^*$(x) $\geq$ $inf_{x\in G}\Lambda^*$(x) und das ist nach Definition 3.6. äquivalent zu: Λ^*(G $\cap$ ri(D_{Λ^*})) $\geq$ Λ^*(G $\cap$ F) $\geq$ Λ^*(G). Auf Grund des Lemmas von Rockafellar reicht es nun, zu zeigen:

$$\Lambda^*(G \cap ri(D_{\Lambda^*})) \leq \Lambda^*(G)$$

Falls G $\cap$ $D_{\Lambda*}$ = $\emptyset$ $\to$ G $\cap$ $ri(D_{\Lambda^*})$ = $\emptyset$ $\to$ Λ^*(G $\cap$ $ri(D_{\Lambda^*})$) = ∞ $\to$ Λ^*(G) = ∞, also nehmen wir an, dass D_{Λ^*} nicht-leer ist. Sei nun z $\in$ $ri(D_{\Lambda^*})$ und y $\in$ G $\cap$ D_{Λ^*}.
1. Fall: D_{Λ^*} besteht aus genau einem Punkt $\to$ y = z.
2. Fall: D_{Λ^*} besteht aus unendlich vielen Punkten. Da es sich um eine konvexe Menge handelt, liegt die Strecke $\overline{yz}$ in D_{Λ^*} und wir finden ein α > 0, sodass αz + (1 - α)y $\in$ G $\cap$ ri(D_{Λ^*}).
Nun folgt mit obiger Definition:

$$\begin{aligned}
\Lambda^*(G \cap ri(D_{\Lambda^*})) &\leq \Lambda^*(\alpha z + (1-\alpha)y) \\
&\leq \lim_{\alpha\to 0} \Lambda^*(\alpha z + (1-\alpha)y) \\
&\leq^{\Lambda^* konvex} \lim_{\alpha\to 0} [\alpha\Lambda^*(z) + (1-\alpha)\Lambda^*(y)] \\
&=^{\text{Stetigkeit von } \Lambda^* \text{ auf } ri(D_{\Lambda^*})} \Lambda^*(y)
\end{aligned}$$

Bilden wir nun das Infimum über alle y $\in$ G $\cap$ D_{Λ^*}, erhalten wir Λ^*(G $\cap$ ri(D_{Λ^*})) $\leq$ Λ^*(G $\cap$ D_{Λ^*}) $\leq$ Λ^*(G) und damit die gewünschte Behauptung.

$\square$

Literatur

[1] den Hollander, Frank: Large deviations
Fields Institute Monographs, 2000

[2] Dembo, Amir; Zeitouni, Ofer: Large deviation techniques and applications
Springer, 2. Auflage 1988

[3] König, Wolfgang: Große Abweichungen
www.mathematik.uni-leipzig.de/koenig/www/GA.ps