KAPITEL 7

Ordnungsstatistiken

Seien $X_1, \dots, X_n$ unabhängige und identisch verteilte (u.i.v.) Zufallsvariablen mit Verteilungsfunktion F. Ordnen wir die Stichprobe $X_1, \dots, X_n$ monoton aufsteigend an, so erhalten wir die sogenannten **Ordnungstatistiken**

$$X_{1:n} \leq X_{2:n} \leq \ldots \leq X_{n:n}.$$

Zum Beispiel ist $X_{1:n} = \min\{X_1, \dots, X_n\}$ und $X_{n:n} = \max\{X_1, \dots, X_n\}$.

7.1. Allgemeine Eigenschaften der Ordnungsstatistiken

Zuerst berechnen wir die Verteilungsfunktion der Ordnungsstatistik $X_{k:n}$.

Satz 7.1.1. Für alle $1 \leq k \leq n$ und $t \in \mathbb{R}$ gilt

$$\mathbb{P}[X_{k:n} \leq t] = \sum_{m=k}^{n} \binom{n}{m} F(t)^m \bar{F}(t)^{n-m}.$$

Beweis. Es sei $L := \sum_{i=1}^{n} \mathbb{1}_{X_i \leq t}$ die Anzahl der Elemente der Stichprobe $X_1, \dots, X_n$, die unterhalb von t liegen. Dann gilt $L \sim \text{Bin}(n, F(t))$ und somit

$$\mathbb{P}[X_{k:n} \leq t] = \mathbb{P}[L \geq k] = \sum_{m=k}^{n} \mathbb{P}[L = m] = \sum_{m=k}^{n} \binom{n}{m} F(t)^m \bar{F}(t)^{n-m}.$$

Im letzten Schritt haben wir die Formel für die Zähldichte einer Binomialverteilung benutzt. □

Sind die u.i.v. Zufallsvariablen $X_1, \dots, X_n$ absolut stetig, so können wir die Dichte der Ordnungsstatitsik $X_{k:n}$ berechnen.

Satz 7.1.2. Es seien $X_1, \dots, X_n$ u.i.v. Zufallsvariablen mit Dichte f und Verteilungsfunktion F. Dann ist für alle $1 \leq k \leq n$ die Dichte von $X_{k:n}$ gegeben durch

$$f_{X_{k:n}}(t) = n\binom{n-1}{k-1} f(t) F(t)^{k-1} \bar{F}(t)^{n-k} = k\binom{n}{k} f(t) F(t)^{k-1} \bar{F}(t)^{n-k}.$$

Beweisidee. Man kann den Satz beweisen, indem man die Formel aus Satz 7.1.1 ableitet (siehe z.B. Skript “Mathematische Statistik”, Satz 1.6.1). Dieser Weg führt zu komplizierten

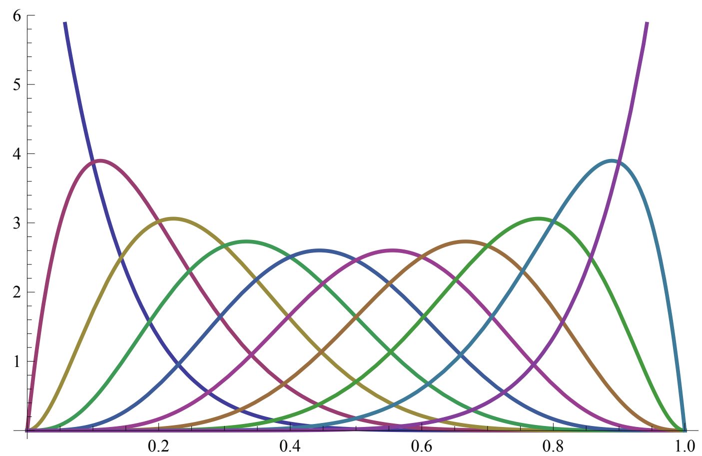

ABBILDUNG 1. Dichten der Ordnungsstatistiken $X_{1:10}, \dots, X_{10:10}$ einer unabhängigen und auf $[0,1]$ gleichverteilten Stichprobe $X_1, \dots, X_{10}$.

Berechnungen. Wir geben hier einen anderen Beweis, der viel eleganter (allerdings nicht ganz streng) ist.

Damit $X_{k:n} = t$ ist, muss Folgendes passieren:

1. Eine der Zufallsvariablen, z.B. X_i, muss den Wert t annehmen. Es gibt n Möglichkeiten, das i auszuwählen. Die "Dichte" des Ereignisses $X_i = t$ ist $f(t)$.

2. Unter den restlichen $n-1$ Zufallsvariablen müssen genau $k-1$ Zufallsvariablen Werte kleiner als t annehmen. Wir haben $\binom{n-1}{k-1}$ Möglichkeiten, die $k-1$ Zufallsvariablen auszuwählen. Die Wahrscheinlichkeit, dass die ausgewählten Zufallsvariablen allesamt kleiner als t sind, ist $F(t)^{k-1}$.

3. Die verbliebenen $n-k$ Zufallsvariablen müssen allesamt größer als t sein. Die Wahrscheinlichkeit davon ist $(1-F(t))^{n-k}$.

Indem wir nun alles ausmultiplizieren, erhalten wir die Behauptung des Satzes. Den kombinatorischen Faktor kann man auch anders berechnen: Zuerst wählen wir aus n Zufallsvariablen k Zufallsvariablen, die $\leq t$ sind (dafür digt es $\binom{n}{k}$ Möglichkeiten), und dann wählen aus diesen k Zufallsvariablen eine Zufallsvariable, der gleich t sein soll (dafür gibt es k Möglichkeiten). □

Beispiel 7.1.3. Seien $X_1, \dots, X_n$ unabhängig und gleichverteilt auf $[0,1]$, d.h. die Dichte von X_i sei $f(t) = \mathbb{1}_{[0,1]}(t)$. Dann ist die Dichte von $X_{k:n}$ gegeben durch

$$f_{X_{k:n}}(t) = \begin{cases} k\binom{n}{k} t^{k-1}(1-t)^{n-k}, & t \in [0,1], \\ 0, & \text{sonst.} \end{cases}$$

Das heißt, $X_{k:n}$ hat eine Beta–Verteilung Beta$(k, n-k+1)$.

Aufgabe 7.1.4. Zeigen Sie, dass $\mathbb{E}X_{k:n} = \frac{k}{n+1}$ für alle $1 \leq k \leq n$.

Aufgabe 7.1.5. Seien $X_1, \dots, X_n$ unabhängige und identisch verteilte Zufallsvariablen mit Dichte f und Verteilungsfunktion F. Bestimmen Sie für $1 \le i < j \le n$ die gemeinsame Dichte $f_{X_{i:n}, X_{j:n}}(t, s)$ der Ordnungsstatistiken $X_{i:n}$ und $X_{j:n}$.

Im nächsten Satz bestimmen wir die gemeinsame Dichte *aller* n Ordnungsstatistiken.

Satz 7.1.6. Seien $X_1, \dots, X_n$ unabhängig und identisch verteilt mit Dichte f. Dann gilt für die gemeinsame Dichte von $X_{1:n}, \dots, X_{n:n}$:

$$f_{X_{1:n},\dots,X_{n:n}}(t_1, \dots, t_n) = \begin{cases} n! \cdot f(t_1) \cdot \ldots \cdot f(t_n), & t_1 < t_2 < \ldots < t_n, \\ 0, & \text{sonst.} \end{cases}$$

Beweisidee. Da die Ordnungsstatistiken per Definition aufsteigend sind, ist die Dichte gleich 0, wenn die Bedingung $t_1 < \ldots < t_n$ nicht erfüllt ist. Sei also die Bedingung $t_1 < \ldots < t_n$ erfüllt. Damit $X_{1:n} = t_1, \dots, X_{n:n} = t_n$ ist, muss eine der Zufallsvariablen (für deren Wahl es n Möglichkeiten gibt) gleich t_1 sein, eine andere (für deren Wahl es $n-1$ Möglichkeiten gibt) gleich t_2, usw. Wir haben also $n!$ Möglichkeiten für die Wahl der Reihenfolge der Variablen. Zum Beispiel tritt für $n = 2$ das Ereignis $\{X_{1:2} = t_1, X_{2:2} = t_2\}$ genau dann ein, wenn entweder $\{X_1 = t_1, X_2 = t_2\}$ oder $\{X_1 = t_2, X_2 = t_1\}$ eintritt, was 2 Möglichkeiten ergibt. Da alle Möglichkeiten sich nur durch Permutationen unterscheiden und somit die gleiche "Dichte" besitzen, betrachten wir nur eine Möglichkeit und multiplizieren dann das Ergebnis mit $n!$. Die einfachste Möglichkeit ist, dass $\{X_{1:n} = t_1, \dots, X_{n:n} = t_n\}$ eintritt. Diesem Ereignis entspricht die "Dichte" $f(t_1) \cdot \ldots \cdot f(t_n)$, da die Zufallsvariablen $X_1, \dots, X_n$ unabhängig sind. Multiplizieren wir nun diese Dichte mit $n!$, so erhalten wir das gewünschte Ergebnis. □

Auch wenn die Zufallsvariablen $X_1, \dots, X_n$ unabhängig sind, sind die Ordnungsstatistiken $X_{1:n}, \dots, X_{n:n}$ im Allgemeinen nicht unabhängig. Es gilt jedoch eine schwächere Eigenschaft, die sogenannte Markov-Eigenschaft.

Satz 7.1.7 (Bedingte Unabhängigkeit von Ordnungsstatistiken)**.** Seien $X_1, \dots, X_n$ unabhängige und identisch verteilte Zufallsvariablen mit Dichte f. Die bedingte Dichte $f(u|x_1, \dots, x_k)$ von $X_{k+1:n}$ gegeben, dass $X_{1:n} = x_1, \dots, X_{k:n} = x_k$, stimmt mit der bedingten Dichte $f(u|x_k)$ von $X_{k+1:n}$ gegeben, dass $X_{k:n} = x_k$, überein.

Bemerkung 7.1.8. Angenommen, die ersten k Ordnungsstatistiken der Stichprobe $X_1, \dots, X_n$ sind bekannt: $X_{1:n} = x_1, \dots, X_{k:n} = x_k$. Wo liegt nun die nächste Ordnungsstatistik $X_{k+1:n}$? Der obige Satz behauptet, dass für die Beantwortung dieser Frage nur der Wert $X_{k:n} = x_k$ relevant ist. Die Werte der vorherigen Ordnungsstatistiken $x_1, \dots, x_{k-1}$ tauchen in der bedingten Verteilung von $X_{k+1:n}$ nicht auf.

Beweis. Sei $f_{1,\ldots,k:n}$ die gemeinsame Dichte von $X_{1:n},\ldots,X_{k:n}$. Dann gilt

$$f(u|x_1,\ldots,x_k)=\frac{f_{1,\ldots,k+1:n}(x_1,\ldots,x_k,u)}{f_{1,\ldots,k:n}(x_1,\ldots,x_k)}. \tag{7.1.1}$$

Sei $f_{k,k+1:n}$ die gemeinsame Dichte von $(X_{k:n},X_{k+1:n})$ und $f_{k:n}$ die Dichte von $X_{k:n}$. Dann gilt

$$f(u|x_k)=\frac{f_{k,k+1:n}(x_k,u)}{f_{k:n}(x_k)}. \tag{7.1.2}$$

Für die gemeinsame Dichte von $X_{1:n},\ldots,X_{k:n}$ erhält man

$$f_{1,\ldots,k:n}(x_1,\ldots,x_k)=f(x_1)\cdot\ldots\cdot f(x_k)\cdot(1-F(x_k))^{n-k}\cdot n\cdot(n-1)\cdot\ldots\cdot(n-k+1),$$

falls $x_1<\ldots<x_k$ und 0 sonst.
Analog erhält man für die gemeinsame Dichte von $X_{1:n},\ldots,X_{k+1:n}$

$$f_{1,\ldots,k+1:n}(x_1,\ldots,x_{k+1})=f(x_1)\cdot\ldots\cdot f(x_{k+1})\cdot(1-F(x_{k+1}))^{n-k-1}\cdot n\cdot(n-1)\cdot\ldots\cdot(n-k),$$

falls $x_1<\ldots<x_{k+1}$ und 0 sonst. Außerdem gilt:

$$f_{k,k+1:n}(x_k,u)=f(x_k)\cdot f(u)\cdot F(x_k)^{k-1}(1-F(u))^{n-k-1}\cdot n(n-1)\binom{n-2}{k-1}$$

bzw.

$$f_{k:n}(x_k)=f(x_k)\cdot F(x_k)^{k-1}(1-F(x_k))^{n-k}\cdot n\binom{n-1}{k-1}.$$

Einsetzen in (7.1.1) bzw. (7.1.2) ergibt

$$(7.1.1)=(7.1.2)=\frac{(n-k)(1-F(u))^{n-k-1}f(u)}{(1-F(x_k))^{n-k}}.$$

Somit stimmen die beiden bedingten Dichten $f(u|x_1,\ldots,x_k)$ und $f(u|x_k)$ überein. □

Aufgabe 7.1.9. Seien $X_1,\ldots,X_n$ unabhängig und exponentialverteilt mit Parameter $\lambda>0$. Welche Verteilung besitzt $X_{k+1:n}-x_k$ gegeben, dass $X_{1:n}=x_1,\ldots,X_{k:n}=x_k$?

7.2. Extreme Ordnungsstatistiken

Es seien $X_1,X_2,\ldots$ unabhängige identisch verteilte Zufallsvariablen. Wir bezeichnen mit $X_{1:n}\le\ldots\le X_{n:n}$ die Ordnungsstatistiken von $X_1,\ldots,X_n$. Außerdem benutzen wir die Notation

$$M_n^{(k)}=X_{n-k+1:n},\qquad k=1,\ldots,n.$$

Somit ist $M_n^{(1)}=X_{n:n}$ das größte Element der Stichprobe, $M_n^{(2)}=X_{n-1:n}$ das zweitgrößte Element und so weiter.

Wir haben die möglichen Grenzwertverteilungen von $M_n^{(1)}=M_n$ für $n\to\infty$ bereits in früheren Kapiteln beschrieben (Extremwertverteilungen). In diesem Abschnitt beschreiben wir die Grenzverteilungen der sogenannten *extremen Ordnungsstatistiken* $M_n^{(k)}$, wobei $k \in \mathbb{N}$ fest und $n \to \infty$.

Zuerst benötigen wir einen Satz, der die Grenzwertverteilung für die Anzahl der Überschreitungen eines hohen Schwellenwerts beschreibt.

Satz 7.2.1. Sei $u_n \in \mathbb{R}$ eine Folge mit $\lim_{n\to\infty} n\bar{F}(u_n) = \tau$, wobei $0 \le \tau < \infty$. Dann gilt:

$$\lim_{n\to\infty} \mathbb{P}\left[\sum_{i=1}^{n} \mathbb{1}_{X_i > u_n} = k\right] = \mathrm{e}^{-\tau}\frac{\tau^k}{k!}, \quad k = 0, 1, \dots.$$

Beweis. Wir bezeichnen mit $S_n := \sum_{i=1}^n \mathbb{1}_{X_i > u_n}$ die Anzahl der Beobachtungen, die oberhalb von u_n liegen. Es gilt $S_n \sim \mathrm{Bin}(n, \bar{F}(u_n))$. Nach Voraussetzung gilt $\lim_{n\to\infty} n\bar{F}(u_n) = \tau$. Da dies gilt, darf der Poissongrenzwertsatz angewendet werden:

$$S_n \xrightarrow[n\to\infty]{d} \mathrm{Poi}(\tau).$$

Mit anderen Worten, $\lim_{n\to\infty} \mathbb{P}[S_n = k] = \mathrm{e}^{-\tau}\frac{\tau^k}{k!}$, für alle $k = 0, 1, \dots$. □

Bemerkung 7.2.2. Der Satz behält seine Gültigkeit für $\tau = +\infty$ (niedriger Schwellenwert, sehr viele Überschreitungen) im folgenden Sinne:

$$\lim_{n\to\infty} \mathbb{P}\left[\sum_{i=1}^{n} \mathbb{1}_{X_i > u_n} = k\right] = 0, \quad k = 0, 1, \dots.$$

Eine $\mathrm{Poi}(+\infty)$–verteilte Zufallsvariable interpretiert man dabei als eine Zufallsvariable, die mit Wahrscheinlichkeit 1 den Wert $+\infty$ annimmt.

Satz 7.2.3. Es gebe Normierungsfolgen $a_n > 0$ und $b_n \in \mathbb{R}$, so dass

$$\frac{M_n^{(1)} - b_n}{a_n} \xrightarrow[n\to\infty]{d} G,$$

wobei G eine Extremwertverteilung ist. Dann gilt für alle $k = 1, 2, \dots$ und $x \in \mathbb{R}$:

$$\lim_{n\to\infty} \mathbb{P}\left[\frac{M_n^{(k)} - b_n}{a_n} \le x\right] = \begin{cases} G(x) \cdot \sum_{s=0}^{k-1} \frac{(-\log G(x))^s}{s!}, & G(x) \neq 0, \\ 0, & G(x) = 0. \end{cases}$$

Beweis. Setzt man $u_n := a_n x + b_n$ und $S_n = \sum_{i=1}^n \mathbb{1}_{X_i > u_n}$, so ergibt sich:

$$\mathbb{P}\left[M_n^{(k)} \le a_n x + b_n\right] = \mathbb{P}[S_n \le k-1] = \sum_{s=0}^{k-1} \mathbb{P}[S_n = s].$$

Mit Satz 7.2.1 folgt:

$$\lim_{n\to\infty} \sum_{s=0}^{k-1} \mathbb{P}[S_n = s] = \sum_{s=0}^{k-1} \mathrm{e}^{-\tau} \frac{\tau^s}{s!} \tag{7.2.1}$$

mit $\tau = \lim_{n\to\infty} n\bar{F}(u_n)$. Es bleibt also noch die Bestimmung von τ. Nach Voraussetzung gilt:

$$\lim_{n\to\infty} F^n(a_n x + b_n) = G(x).$$

Für $G(x) \neq 0$ ist dies äquivalent zu

$$\lim_{n\to\infty} n \log(1 - \bar{F}(u_n)) = \log G(x).$$

Unter Verwendung der Taylorentwicklung des Logarithmus $\log(1-x) = -x + o(x)$, $x \to 0$, folgt

$$\tau = \lim_{n\to\infty} nF(u_n) = -\log G(x)$$

Einsetzen von τ in (7.2.1) liefert die Behauptung. Für $G(x) = 0$ ist der Beweis analog, liefert allerdings $\tau = +\infty$. $\square$

Im Folgenden bescheiben wir die Grenzwertverteilung des Vektors $(M_n^{(1)}, \dots, M_n^{(r)})$, mit r fest und $n \to \infty$. Dazu benötigen wir eine Verallgemeinerung von Satz 7.2.1, die r verschiedene Folgen von Schwellenwerten zulässt. Betrachte die Schwellenwerte $u_n^{(1)} \geq u_n^{(2)} \geq \dots \geq u_n^{(r)}$ mit

$$\lim_{n\to\infty} n\bar{F}(u_n^{(1)}) = \tau_1, \quad \lim_{n\to\infty} n\bar{F}(u_n^{(2)}) = \tau_1 + \tau_2, \quad \dots, \quad \lim_{n\to\infty} n\bar{F}(u_n^{(r)}) = \tau_1 + \dots + \tau_r,$$

wobei $\tau_1, \dots, \tau_r \in [0, \infty)$. Sei

$$S_n^{(l)} := \sum_{i=1}^{n} \mathbb{1}_{X_i > u_n^{(l)}}$$

die Anzahl der Überschreitungen des Schwellenwerts $u_n^{(l)}$, $1 \leq l \leq r$.

Satz 7.2.4. Für alle $k_1, k_2, \dots, k_r \in \{0, 1, 2, \dots\}$ gilt

$$\lim_{n\to\infty} \mathbb{P}[S_n^{(1)} = k_1, S_n^{(2)} = k_1 + k_2, \dots, S_n^{(r)} = k_1 + \dots + k_r] = \mathrm{e}^{-(\tau_1 + \dots + \tau_r)} \frac{\tau_1^{k_1} \tau_2^{k_2} \dots \tau_r^{k_r}}{k_1! k_2! \dots k_r!}.$$

Bemerkung 7.2.5. Mit anderen Worten, es besteht die folgende Verteilungskonvergenz

$$(S_n^{(1)}, S_n^{(2)} - S_n^{(1)}, \dots, S_n^{(r)} - S_n^{(r-1)}) \xrightarrow[n\to\infty]{d} (\mathrm{Poi}(\tau_1), \dots, \mathrm{Poi}(\tau_r)),$$

wobei die Komponenten der Grenzwertverteilung unabhängig sind.

Beweis. Definiere $p_{n,l} = \bar{F}(u_n^{(l)}) - \bar{F}(u_n^{(l-1)})$ für $1 \leq l \leq r$, wobei $\bar{F}(u_n^{(0)})$ als 0 interpretiert wird. Setzt man

$$P_n(k_1, \dots, k_n) = \mathbb{P}[S_n^{(1)} = k_1, S_n^{(2)} = k_1 + k_2, \dots, S_n^{(r)} = k_1 + \dots + k_r],$$

dann gilt

$$P_n(k_1,\dots,k_n) = \binom{n}{k_1} p_{n,1}^{k_1} \cdot \binom{n-k_1}{k_2} p_{n,2}^{k_2} \cdot \ldots \cdot \binom{n-\sum_{i=1}^{r-1} k_i}{k_r} p_{n,r}^{k_r} \cdot \cdot (1-p_{n,1}-\ldots-p_{n,r})^{n-\sum_{i=1}^{r} k_i}. \tag{7.2.2}$$

Damit kann der Beweis faktorweise durchgeführt werden. Für den ersten Faktor gilt

$$\binom{n}{k_1} p_{n,1}^{k_1} = \frac{n(n-1)\dots(n-k_1+1)}{k_1!} \cdot \bar{F}(u_n^{(1)})^{k_1} \sim \frac{n^{k_1}\bar{F}(u_n^{(1)})^{k_1}}{k_1!} \underset{n\to\infty}{\longrightarrow} \frac{\tau_1^{k_1}}{k_1!}.$$

Ähnlich gilt für den zweiten Faktor

$$\binom{n-k_1}{k_2} p_{n,2}^{k_2} \sim \frac{n^{k_2}(\bar{F}(u_n^{(2)})-\bar{F}(u_n^{(1)}))^{k_2}}{k_2!} \underset{n\to\infty}{\longrightarrow} \frac{\tau_2^{k_2}}{k_2!},$$

was sich nun analog für sämtliche Faktoren, außer dem letzten, fortführen lässt. Für den letzten Faktor gilt schließlich:

$$(1-p_{n,1}-\ldots-p_{n,r})^{n-k_1-\ldots-k_r} \underset{n\to\infty}{\longrightarrow} \mathrm{e}^{-(\tau_1+\ldots+\tau_r)},$$

denn es gilt:

$$(n-k_1-\ldots-k_r)(p_{n,1}+\ldots+p_{n,r}) = (n-k_1-\ldots-k_r)\bar{F}(u_n^{(r)}) \underset{n\to\infty}{\longrightarrow} \tau_1+\ldots+\tau_r.$$

Setzt man alles zusammen, so erhält man die Behauptung. □

Nun können wir die gemeinsame Verteilung der oberen r extremen Ordnungsstatistiken $(M_n^{(1)},\dots,M_n^{(r)})$ beschreiben. Es gilt nämlich

$$\mathbb{P}[M_n^{(1)} \le u_n^{(1)},\dots,M_n^{(r)} \le u_n^{(r)}] = \mathbb{P}[S_n^{(1)}=0, S_n^{(2)} \le 1, S_n^{(3)} \le 2\dots, S_n^{(r)} \le r-1].$$

Der Grenzwert der rechten Seite kann mit dem obigen Satz berechnet werden, was zu einer langen Formel führt. Im nächsten Satz betrachten wir den Spezialfall $r=2$, d.h. wir beschreiben die gemeinsame Grenzwertverteilung der zwei größten Werte der Stichprobe.

Satz 7.2.6. Es gebe Normierungsfolgen $a_n>0$ und $b_n\in\mathbb{R}$, so dass

$$\frac{M_n^{(1)}-b_n}{a_n} \underset{n\to\infty}{\overset{d}{\longrightarrow}} G,$$

wobei G eine Extremwertverteilung ist. Dann gilt für alle $x_1>x_2$:

$$\lim_{n\to\infty} \mathbb{P}\left[\frac{M_n^{(1)}-b_n}{a_n} \le x_1, \frac{M_n^{(2)}-b_n}{a_n} \le x_2\right] = G(x_2)(\log G(x_1) - \log G(x_2) + 1),$$

falls $G(x_2)\neq 0$ und 0 sonst.

Beweis. Setze $u_n^{(1)} := a_n x_1 + b_n$ bzw. $u_n^{(2)} := a_n x_2 + b_n$. Es ergibt sich aus

$$\lim_{n\to\infty} F^n(a_n x + b_n) = G(x) \Rightarrow \lim_{n\to\infty} n\bar{F}(a_n x + b_n) = -\log G(x),$$

dass

$$\tau_1 := \lim_{n\to\infty} n\bar{F}(u_n^{(1)}) = -\log G(x_1),$$
$$\tau_2 := \lim_{n\to\infty} n\bar{F}(u_n^{(2)}) = -\log G(x_2).$$

Es gilt:

$$\begin{aligned}\mathbb{P}[M_n^{(1)} \le u_n^{(1)}, M_n^{(2)} \le u_n^{(2)}] &= \mathbb{P}[S_n^{(1)} = 0, S_n^{(2)} \le 1] \\ &= \mathbb{P}[S_n^{(1)} = 0, S_n^{(2)} = 0] + \mathbb{P}[S_n^{(1)} = 0, S_n^{(2)} = 1].\end{aligned}$$

Mit Satz 7.2.4 folgt dann

$$\lim_{n\to\infty} \mathbb{P}[M_n^{(1)} \le u_n^{(1)}, M_n^{(2)} \le u_n^{(2)}] = \mathrm{e}^{-\tau_2} + (\tau_2 - \tau_1)\mathrm{e}^{-\tau_2}.$$

Setzt man die Werte von τ_1 und τ_2 ein, so erhält man die Behauptung. □

Bemerkung 7.2.7. Für ein $r > 2$ ist die Formel für die Verteilungsfunktion der Grenzwertverteilung von $(M_n^{(1)}, \ldots, M_n^{(r)})$ sehr kompliziert. Allerdings ist die Formel für die Dichte dieser Verteilung sehr einfach; wir werden diese später unter Verwendung von Poisson–Prozessen herleiten.

7.3. Darstellungen der Ordnungsstatistiken

Eine Zufallsvariable Z heißt exponentialverteilt mit Parameter $\lambda > 0$, falls

$$\mathbb{P}[Z > t] = \mathrm{e}^{-\lambda t} \text{ für } t \ge 0.$$

Die Ordnungsstatistiken der exponentialverteilten Zufallsvariablen besitzen eine besonders schöne Darstellung.

Satz 7.3.1. Seien $Z_1, \ldots, Z_n$ unabhängige und mit Parameter $\lambda = 1$ exponentialverteilte Zufallsvariablen. Dann gilt

$$(Z_{1:n}, Z_{2:n}, \ldots, Z_{n:n}) \stackrel{d}{=} \left(\frac{\nu_1}{n}, \frac{\nu_1}{n} + \frac{\nu_2}{n-1}, \ldots, \frac{\nu_1}{n} + \frac{\nu_2}{n-1} + \ldots + \frac{\nu_n}{1}\right),$$

wobei $\nu_1, \ldots, \nu_n$ unabhängig und exponentialverteilt mit Parameter $\lambda = 1$ sind.

Bemerkung 7.3.2. Die Abstände $Z_{1:n}, Z_{2:n} - Z_{1:n}, \ldots, Z_{n:n} - Z_{n-1:n}$ sind somit unabhängig und (nicht identisch) exponentialverteilt:

$$Z_{k:n} - Z_{k-1:n} \stackrel{d}{=} \frac{\nu_k}{n-k+1} \sim \mathrm{Exp}(n-k+1).$$

Beweis von Satz 7.3.1. Die Dichte von $(Z_{1:n}, \ldots, Z_{n:n})$ ist gegeben durch

$$f(x_1, \ldots, x_n) = \begin{cases} n! \cdot \mathrm{e}^{-x_1} \cdot \ldots \cdot \mathrm{e}^{-x_n}, & 0 < x_1 < \ldots < x_n, \\ 0, & \text{sonst.} \end{cases}$$

Wegen der Unabhängigkeit ist die Dichte von $(\nu_1, \ldots, \nu_n)$ gegeben durch

$$g(y_1, \ldots, y_n) = \begin{cases} \mathrm{e}^{-y_1} \cdot \ldots \cdot \mathrm{e}^{-y_n}, & y_i > 0 \text{ für alle } i \in \{1, \ldots, n\}, \\ 0, & \text{sonst.} \end{cases}$$

Wir betrachten die lineare Transformation

$$T : (y_1, \ldots, y_n) \mapsto \left(\frac{y_1}{n}, \frac{y_1}{n} + \frac{y_2}{n-1}, \ldots, \frac{y_1}{n} + \frac{y_2}{n-1} + \ldots + \frac{y_n}{1} \right) =: (z_1, z_2, \ldots, z_n).$$

Die Jacobi–Determinante dieser Transformation ist

$$J := \begin{vmatrix} \frac{1}{n} & 0 & 0 & \cdots & 0 \\ \frac{1}{n} & \frac{1}{n-1} & 0 & \cdots & 0 \\ \vdots & \vdots & \ddots & \ddots & \vdots \\ \frac{1}{n} & \frac{1}{n-1} & \cdots & \frac{1}{2} & 0 \\ \frac{1}{n} & \frac{1}{n-1} & \cdots & \frac{1}{2} & 1 \end{vmatrix} = \frac{1}{n} \cdot \frac{1}{n-1} \cdot \ldots \cdot \frac{1}{2} \cdot 1 = \frac{1}{n!}.$$

Mit der Transformationsformel können wir die Dichte von $T(\nu_1, \ldots, \nu_n)$ berechnen:

$$h(z_1, \ldots, z_n) = n! g(y_1, \ldots, y_n) = \begin{cases} n! \mathrm{e}^{-(y_1 + \ldots + y_n)} = n! \mathrm{e}^{-(z_1 + \ldots + z_n)}, & 0 < z_1 < \ldots < z_n, \\ 0, & \text{sonst.} \end{cases}$$

Bei obiger Transformation wurde die Dichte von $(\nu_1, \ldots, \nu_n)$ durch die Jacobi–Determinante geteilt, um zur Dichte von $T(\nu_1, \ldots, \nu_n)$ zu gelangen. Man sieht, dass die Dichte f von $(Z_{1:n}, \ldots, Z_{n:n})$ und die Dichte h von $T(\nu_1, \ldots, \nu_n)$ übereinstimmen, was die Behauptung impliziert. □

Bemerkung 7.3.3. Es seien $Z_1, Z_2, \ldots \sim \operatorname{Exp}(1)$ unabhängig. Für das Maximum $M_n = \max\{Z_1, \ldots, Z_n\}$ gilt mit Satz 7.3.1

$$\mathbb{E} M_n = \mathbb{E} Z_{n:n} = \sum_{k=1}^{n} \frac{1}{k} = \log n + \gamma + o(1), \quad n \to \infty,$$

wobei $\gamma \approx 0.5772$ die Euler–Konstante ist. Auf der anderen Seite gilt aber auch

$$M_n - \log n \xrightarrow[n \to \infty]{d} G,$$

wobei G eine Gumbel–verteilte Zufallsvariable ist. Wendet man nun auf beide Seiten den Erwartungswert an (die Begründung lassen wir weg), so erhält man eine Formel für der Erwartungswert der Gumbel–Verteilung:

$$\mathbb{E} G = \gamma.$$

Analog lässt sich zeigen, dass $\operatorname{Var} G = \sum_{k=1}^{\infty} \frac{1}{k^2} = \frac{\pi^2}{6}$.

Aufgabe 7.3.4. Berechnen Sie mit Satz 7.3.1 $\mathbb{E} Z_{k:n}$, $\operatorname{Var} Z_{k:n}$ und $\operatorname{Cov}(Z_{k:n}, Z_{l:n})$.

Die Ordnungsstatistiken der Zufallsvariablen, die gemäß einer Gleichverteilung verteilt sind, besitzen ebenfalls eine schöne Darstellung.

Satz 7.3.5. Die Zufallsvariablen $U_1, \ldots, U_n$ seien unabhängig und gleichverteilt auf dem Intervall $[0,1]$. Dann gilt:

$$(U_{1:n}, U_{2:n}, \ldots, U_{n:n}) \overset{d}{=} \left(\frac{S_1}{S_{n+1}}, \frac{S_2}{S_{n+1}}, \ldots, \frac{S_n}{S_{n+1}} \right),$$

wobei $S_k = \nu_1 + \ldots + \nu_k$ und $\nu_1, \ldots, \nu_{n+1}$ unabhängige, mit Parameter 1 exponentialverteilte Zufallsvariablen sind.

Bemerkung 7.3.6. Die Folge $S_1 < S_2 < \ldots$ heißt der **Poisson–Prozess** auf $(0, \infty)$ mit Intensität 1. Intuitiv kann man die Aussage des Satzes so verstehen, dass die ersten n Punkte $S_1, \ldots, S_n$ des Poisson–Prozesses auf dem Intervall $[0, S_{n+1}]$ gleichverteilt sind.

Beweis. Die Dichte von $(U_{1:n}, \ldots, U_{n:n})$ ist gegeben durch:

$$f(x_1, \ldots, x_n) = \begin{cases} n!, & 0 < x_1 < x_2 < \ldots < x_n < 1, \\ 0, & \text{sonst.} \end{cases}$$

Die Dichte von $(\nu_1, \ldots, \nu_{n+1})$ ist wegen der Unabhängigkeit gegeben durch:

$$g(y_1, \ldots, y_{n+1}) = \begin{cases} \mathrm{e}^{-y_1} \cdot \ldots \cdot \mathrm{e}^{-y_{n+1}}, & y_i > 0 \text{ für alle } 1 \le i \le n+1, \\ 0, & \text{sonst.} \end{cases}$$

Wir betrachten die lineare Transformation

$$T : (y_1, \ldots, y_{n+1}) \mapsto (y_1, y_1 + y_2, \ldots, y_1 + y_2 + \ldots + y_{n+1}) =: (z_1, z_2, \ldots, z_{n+1}).$$

Die Jacobi–Determinante von T ist

$$J_T := \begin{vmatrix} 1 & 0 & 0 & \cdots & 0 \\ 1 & 1 & 0 & \ldots & 0 \\ \vdots & \vdots & \ddots & \ddots & \vdots \\ 1 & 1 & \cdots & 1 & 0 \\ 1 & 1 & \cdots & 1 & 1 \end{vmatrix} = 1.$$

Mit der Transformationsformel errechnen wir die Dichte von $(S_1, S_2, \ldots, S_{n+1})$:

$$h(z_1, \ldots, z_{n+1}) = \begin{cases} \mathrm{e}^{-(y_1 + \ldots + y_{n+1})} = \mathrm{e}^{-z_{n+1}}, & 0 < z_1 < z_2 < \ldots < z_{n+1}, \\ 0, & \text{sonst.} \end{cases}$$

Nun betrachten wir eine weitere (diesmal nichtlineare) Transformation

$$S : (z_1, \ldots, z_{n+1}) \mapsto \left(\frac{z_1}{z_{n+1}}, \frac{z_2}{z_{n+1}}, \ldots, \frac{z_n}{z_{n+1}}, z_{n+1} \right) =: (w_1, w_2, \ldots, w_n, w_{n+1}).$$

Die Jacobi–Determinante von S ist

$$J_S(z_1,\dots,z_{n+1}) = \begin{vmatrix} \frac{1}{z_{n+1}} & 0 & 0 & \cdots & 0 \\ 0 & \frac{1}{z_{n+1}} & 0 & \dots & 0 \\ \vdots & \vdots & \ddots & \ddots & \vdots \\ 0 & 0 & \cdots & \frac{1}{z_{n+1}} & 0 \\ \frac{-z_1}{z_{n+1}^2} & \frac{-z_2}{z_{n+1}^2} & \cdots & \frac{-z_n}{z_{n+1}^2} & 1 \end{vmatrix} = \frac{1}{z_{n+1}^n} = \frac{1}{w_{n+1}^n}.$$

Die Dichte von $(\frac{S_1}{S_{n+1}}, \frac{S_2}{S_{n+1}}, \dots, \frac{S_n}{S_{n+1}}, S_{n+1})$ kann nun mit der Transformationsformel wie folgt berechnet werden:

$$p(w_1,\dots,w_{n+1}) = \begin{cases} \mathrm{e}^{-z_{n+1}} z_{n+1}^n = \mathrm{e}^{-w_{n+1}} w_{n+1}^n, & 0 < w_1 < w_2 < \ldots < w_n < 1, w_{n+1} > 0, \\ 0, & \text{sonst.} \end{cases}$$

Die Dichte von $(\frac{S_1}{S_{n+1}}, \frac{S_2}{S_{n+1}}, \dots, \frac{S_n}{S_{n+1}})$ ist eine Marginaldichte von p:

$$r(w_1,\dots,w_n) = \begin{cases} \int_{\mathbb{R}} p(w_1,\dots,w_{n+1}) \mathrm{d}w_{n+1} = n!, & 0 < w_1 < w_2 < \ldots < w_n < 1, \\ 0, & \text{sonst.} \end{cases}$$

Man sieht, dass die Dichte f von $(U_{1:n},\dots,U_{n:n})$ und die Dichte r von $(\frac{S_1}{S_{n+1}}, \frac{S_2}{S_{n+1}}, \dots, \frac{S_n}{S_{n+1}})$ übereinstimmen, woraus die Behauptung folgt. □

Aufgabe 7.3.7. Es seien $U_1,\dots,U_n$ unabhängige, auf dem Intervall $[0,1]$ gleichverteilte Zufallsvariablen. Zeigen Sie, dass für jedes $k \in \mathbb{N}$ die folgende Verteilungskonvergenz gilt:

$$(nU_{1:n},\dots,nU_{k:n}) \xrightarrow[n\to\infty]{d} (\nu_1, \nu_1+\nu_2, \dots, \nu_1+\ldots+\nu_k),$$

wobei $\nu_1,\nu_2,\dots$ unabhängige Zufallsvariablen mit $\mathbb{P}[\nu_i > t] = \mathrm{e}^{-t}$, $t > 0$, sind (Standardexponentialverteilung). Zeigen Sie auch, dass

$$(n(1-U_{n:n}),\dots,n(1-U_{n-k+1:n})) \xrightarrow[n\to\infty]{d} (\nu_1, \nu_1+\nu_2, \dots, \nu_1+\ldots+\nu_k).$$