5.2 Multinomiale Wahrscheinlichkeiten

Wir betrachten ein zufälliges Experiment mit den Ausgängen $A_1, A_2, \ldots, A_l$. Wir setzen

$$p_i = P(A_i), \qquad \sum_{i=1}^{l} p_i = 1.$$

Ein Beispiel ist das folgende Experiment:
Es sei ein Behälter mit k Kugeln in l verschiedenen Farben gegeben, wobei k_i Kugeln die Farbe i ($i = 1, \ldots, l$) besitzen,

$$\sum_{i=1}^{l} k_i = k.$$

Es soll die Wahrscheinlichkeit untersucht werden, mit der eine Kugel einer bestimmten Farbe aus dem Behälter entnommen wird:

$$P(\text{Kugel der Farbe } i) = p_i = \tfrac{k_i}{k}.$$

Das Experiment soll nun n–mal wiederholt werden.

$B_{n_1,n_2,\ldots,n_l}$: das Ereignis, daß die Ereignisse A_1 n_1–mal, A_2 n_2–mal, ..., und A_l n_l–mal eintreten.

$$P(B_{n_1,k_2,\ldots,n_l}) = \frac{n!}{n_1! \cdot n_2! \cdot \ldots \cdot n_l!} \cdot p_1^{n_1} \cdot p_2^{n_2} \cdot \ldots \cdot p_l^{n_l}.$$

Derartige Wahrscheinlichkeiten bezeichnen wir auch als multinomiale Wahrscheinlichkeiten (oder polynomiale Wktn.)

5.3 POISSON–Wahrscheinlichkeiten

Beispiele, bei denen POISSON–Wahrscheinlichkeiten auftreten, sind

- die Anzahl von Verkehrsunfällen in einem Ort in einem bestimmten Zeitintervall,
- die Ankünfte von Kunden an einem Schalter oder
- der radioaktive Zerfall von α–Teilchen.
- In einer Telefonzentrale wird ermittelt, wieviel Anrufe in einer bestimmten Zeiteinheit ankommen.

Elementarereignisse sind hier also zufällige Anzahlen.

$$\Omega = \{\omega_1, \omega_2, \ldots, \omega_n, \ldots\} = \{„0“, „1“, \ldots, „n“, \ldots\}.$$

Das Ereignis ω_i ist z.B. das Ereignis, daß in einer Zeiteinheit genau i Anrufe eintreffen. Die Wahrscheinlichkeit dieses Elementarereignisses ist gegeben durch:

$$P(\omega_i) = \frac{\lambda^i}{i!} e^{-\lambda}.$$

λ ist dabei ein noch unbestimmter Parameter. Er kann als mittlere Rate aufgefaßt werden. Diese Wahrscheinlichkeit bezeichnen wir als <u>POISSON–Wahrscheinlichkeit</u>. Für die Wahrscheinlichkeit von Ω gilt dann:

$$\begin{aligned} P(\Omega) &= \sum_{i=0}^{\infty} P(\omega_i) = \sum_{i=0}^{\infty} \frac{\lambda^i}{i!} e^{-\lambda} \\ &= e^{-\lambda} \underbrace{\sum_{i=0}^{\infty} \frac{\lambda^i}{i!}}_{=e^{\lambda}} = 1 \end{aligned}$$

Wir werden später sehen, daß diese Verteilung “natürlich” ist.

Kapitel 2

Zufallsvariablen

Contents

1 Grundbegriffe 100

2 Eigenschaften der Verteilungsfunktion . 118

3 Diskrete zufällige Variablen 122

4 Charakteristika von Verteilungsfunktionen 142

5 Die Exponentialverteilung 175

6 Die Normalverteilung 204

7 Transformation von Zufallsvariablen . . 219

8 Mehrdimensionale Zufallsvariablen 229

1 Grundbegriffe

Def. 2.1 *Es seien* $(\Omega_1, \mathcal{E}_1, P_1)$ *und* $(\Omega_2, \mathcal{E}_2, P_2)$ *Wahrscheinlichkeitsräume. Eine Abbildung*

$$X\colon \Omega_1 \longrightarrow \Omega_2$$

heißt $\underline{\mathcal{E}_1\text{–}\mathcal{E}_2\text{–meßbar}}$*, falls für alle Ereignisse* $A \in \mathcal{E}_2$ *gilt:*

$$X^{-1}(A) = \{\omega \in \Omega_1\colon X(\omega) \in A\} \in \mathcal{E}_1.$$

Bem.: Oftmals wird die Menge $\mathcal{B}^1$ der BOREL–Mengen als Ereignisfeld $\mathcal{E}_2$ betrachtet.

Def. 2.2 *Es sei* $(\Omega, \mathcal{E}, P)$ *ein Wahrscheinlichkeitsraum. Eine* $\mathcal{E}$*–*$\mathcal{B}^1$*–meßbare Abbildung* X *von* Ω *in* $\mathbb{R}$ *heißt (reellwertige)*
oder Zufallsgröße.

Bem.: $(\mathbb{R}, \mathcal{B}^1, P')$ bildet hier den zweiten Wahrscheinlichkeitsraum, wobei P' eine Abbildung von $\mathcal{B}^1$ in $\mathbb{R}$ ist, die den KOLMOGOROFF–Axiomen genügt.

Wir betrachten den Wahrscheinlichkeitsraum $(\Omega, \mathcal{E}, P)$. $X: \Omega \longrightarrow \mathbb{R}$. sei eine zufällige (reellwertige) Variable.

Den zweiten Wahrscheinlichkeitsraum bezeichnen wir mit $(\mathbb{R}, \mathcal{B}^1, P_X)$. Es sei $B \in \mathcal{B}^1$ ein zufälliges Ereignis, für das gilt:

$$B =]-\infty, x[,$$

wobei x eine beliebige, fest gewählte reelle Zahl ist. Mit $\{X < x\}$ bezeichnen wir das zufällige Ereignis, für das gilt:

$$\{X < x\} := \{\omega \in \Omega \colon X(\omega) < x\}.$$

Dann gilt für die Wahrscheinlichkeit dieses Ereignisses:

$$\begin{aligned} P(X < x) &= P(\{\omega\colon X(\omega) < x\}) = P(\{\omega\colon X(\omega) \in B\}) \\ &= P(X^{-1}(B)) =: P_X(B) \end{aligned}$$

Für alle zufälligen Ereignisse $B \in \mathcal{B}^1$ bezeichnen wir also: $\underline{P_X(B) := P(X^{-1}(B)).}$

Def. 2.3 *Es sei* $X : \Omega \longrightarrow \mathbb{R}$ *eine zufällige Variable,* $(\Omega, \mathcal{E}, P)$ *und* $(\mathbb{R}, \mathcal{B}^1, P_X)$ *seien Wahrscheinlichkeitsräume. Dann heißt die Funktion*

$$F_X(x) := P(X < x) = P_X\left(]-\infty, x[\right)$$

Verteilungsfunktion von X.

Bem.: Der Einfachheit halber werden wir die Funktion F_X einfach nur mit F bezeichnen.

Bem.: Manchmal wird die Verteilungsfunktion auch durch

$$F_X(x) = P(X \leq x)$$

definiert (bei SAS z.B.)

Diskrete Zufallsvariablen

Eine diskrete Zufallsgröße X nimmt höchstens abzählbar viele verschiedene Werte mit positiver Wahrscheinlichkeit an. Das heißt, X ist eine Abbildung der folgenden Form:

$$X\colon \Omega \longrightarrow \{x_i\colon i \in \mathbb{N}\} =: W \subset \mathbb{R}.$$

Wir notieren das in der Form:

$$X : \begin{pmatrix} x_1 & x_2 & \ldots & x_n & \ldots \\ \\ p_1 & p_2 & \ldots & p_n & \ldots \end{pmatrix}$$

Dabei sind die $x_i \in \mathbb{R}$ die Werte, die die Zufallsgröße annehmen kann. Die p_i sind die Wahrscheinlichkeiten, mit denen diese Werte angenommen werden. Es gilt:

$$p_i \geq 0, \qquad \sum_{i=1}^{\infty} p_i = 1, \qquad p_i = P(X = x_i).$$

Wenn wir Mengen A_i definieren durch

$$A_i := \{\omega\colon X(\omega) = x_i\}, \ \ \forall i \in \mathbb{N},$$

so gilt offenbar:

$$A_i \cap A_j = \emptyset, \qquad \forall i, j \in \mathbb{N}, i \neq j.$$

Allgemein gilt dann:

$$P(X = x) = \begin{cases} p_i, & \text{falls } x = x_i \\ 0, & \text{falls } x \neq x_i \end{cases} \qquad \forall x_i \in W,\ i \in \mathbb{N}.$$

Das bedeutet für die Verteilungsfunktion:

$$\begin{aligned} F(x) &= P(X < x) \\ &= P\left(\bigcup_{i:\ x_i < x} A_i\right) \\ &= \sum_{i:\ x_i < x} P(A_i) \\ &= \sum_{i:\ x_i < x} p_i \end{aligned}$$

D.h.: Eine diskrete Zufallsgröße, die die Werte $\{x_i \colon i \in \mathbb{N}\}$ annimmt, wobei $x_1 < x_2 < x_3 < \ldots$ gilt, hat die folgende Verteilungsfunktion:

$$\boxed{F(x) = \begin{cases} 0, & \text{falls } x \leq x_1 \\ \sum\limits_{i:\ x_i < x} p_i, & \text{falls } x_1 < x \end{cases}}$$

Betrachten wir eine Menge $B \in \mathcal{B}^1$, so können wir feststellen:

$$P_X(B) = P(\{\omega\colon X(\omega) \in B\}) = \sum_{i\colon x_i \in B} p_i.$$

Bsp. 2.1 *Es sei*

$$X: \quad \begin{pmatrix} x_1 & x_2 & \ldots & x_n \\ \\ \frac{1}{n} & \frac{1}{n} & \ldots & \frac{1}{n} \end{pmatrix}$$

Die Zufallsvariable X heißt diskret gleichverteilt auf der Menge $\{x_1, \ldots, x_n\}$.

Bsp. 2.2 *Sei X eine diskrete Zufallsgröße,*

$$X: \quad \begin{pmatrix} 0 & 1 & \ldots & n \\ \\ p_0 & p_1 & \ldots & p_n \end{pmatrix}$$

mit

$$P(X = i) = p_i = \binom{n}{i} p^i \cdot (1-p)^{n-i} > 0, \qquad \textit{mit } 0 < p < 1.$$

Bez. 1 *Die Zufallsvariable X heißt binomialverteilt, bez.: $X \sim B(p,n)$ oder $X \sim Bi(p,n)$.*

Wir haben oben gesehen, daß

$$\sum_{i=0}^{n} p_i = \sum_{i=0}^{n} \binom{n}{i} p^i (1-p)^{n-i} = (p+1-p)^n = 1.$$

Bsp. 2.3 *Es sei X eine diskrete Zufallsgröße,*

$$X: \quad \begin{pmatrix} 0 & 1 & \ldots & n & \ldots \\ \\ p_0 & p_1 & \ldots & p_n & \ldots \end{pmatrix}$$

mit

$$P(X=n) = p_n = \frac{\lambda^n}{n!} e^{-\lambda}, \quad \lambda > 0.$$

Bez. 2 *Die Zufallsvariable X heißt* POISSON*–verteilt, bez.: $X \sim Poi(\lambda)$.*

Wir haben oben gesehen, daß

$$\sum_{n=0}^{\infty} p_n = \sum_{n=0}^{\infty} \frac{\lambda^n}{n!} e^{-\lambda} = e^{-\lambda} \underbrace{\sum_{n=0}^{\infty} \frac{\lambda^n}{n!}}_{=e^{\lambda}} = 1$$

Stetige Zufallsvariablen

Def. 2.4 *Eine Funktion $f : \mathbb{R} \longrightarrow \mathbb{R}$ heißt* *Dichtefunktion, falls sie die folgenden Eigenschaften hat:*

1. *Für alle $x \in \mathbb{R}$ gilt: $f(x) \geq 0$.*
2. *Es gilt: $\int\limits_{\mathbb{R}} f(x)\,dx = 1$.*

Def. 2.5 *Eine zufällige Variable X heißt* *stetig, falls eine Dichtefunktion f existiert, so daß gilt:*

$$P(X < x) = F(x) = \int\limits_{-\infty}^{x} f(t)\,dt.$$

Falls die Funktion f stetig ist, gilt: $F'(x) = f(x)$.

Bem.: Für die Wahrscheinlichkeit $P(X = x)$ gilt

$$P(X = x) = \int\limits_{x}^{x} f(t)\,dt = 0,$$

sogar wenn X den Wert x tatsächlich annehmen kann! Das heißt jedoch nichts anderes, als daß gilt:

$$P(X \leq x) = P(X < x).$$

Außerdem gilt:

$$P(a \le X \le b) = \int\limits_a^b f(t)\,dt.$$

Veranschaulichung: Es sei X eine stetige Zufallsgröße. Wir teilen den Wertebereich von X in Intervalle I_j ein und beobachten für jeden der Versuche X_i, in welches der Intervalle I_j der Wert X_i ($i = 1, \ldots, n$) fällt. Es sei $n_j = \#\{X_i \in I_j\}$. Die Länge eines Intervalls I_j bezeichnen wir mit $\Delta(I_j) = \Delta_j$. Desweiteren sei $\Delta_0 = \max\limits_j\{\Delta_j\}$. Wir definieren nun folgende Funktion:

$$f_{emp.}(x) = \frac{\frac{n_j}{n}}{\Delta_j},\ \forall x \in I_j.$$

Dann gilt:

$$f(x) = \lim_{\substack{n\to\infty \\ \Delta_0 \to 0}} f_{emp.}(x).$$

`Dichtefunktion_allg.sas`

Bsp. 2.4 *Es sei die Zufallsvariable X auf dem Intervall $[0, 1[$ definiert mit der Verteilungsfunktion*

$$F(x) = \begin{cases} 0, & \textit{falls } x < 0 \\ x, & \textit{falls } 0 \leq x < 1 \\ 1, & \textit{falls } x \geq 1 \end{cases}.$$

Bez. 3 *Die Zufallsvariable X heißt auf dem Intervall*
$[0, 1[$ gleichverteilt,
bez. $X \sim R(0, 1)$ oder $X \sim U(0, 1)$.

Die Dichtefunktion ist die Funktion f;

$$f(x) = \begin{cases} 0, & \text{falls } x < 0 \\ 1, & \text{falls } 0 \leq x < 1 \\ 0, & \text{falls } x \geq 1 \end{cases}.$$

Ist X gleichverteilt auf dem Intervall $[a, b[$, $X \sim R(a, b)$, so hat X die Dichtefunktion:

$$f(x) = \begin{cases} 0, & \text{falls } x < a \\ \frac{1}{b-a}, & \text{falls } a \leq x < b \\ 0, & \text{falls } x \geq b \end{cases} \cdot$$

Für $0 \leq a < b < 1$ gilt:

$$\begin{aligned} P(\{\omega\colon X(\omega) \in [a, b]) &= P(a \leq X \leq b) \\ &= \int\limits_a^b f(x)\,dx = \frac{1}{b-a} \int_a^b dx = 1 \end{aligned}$$

Bsp. 2.5 *Die Zufallsvariable X habe die Verteilungsfunktion*

$$F(x) = \begin{cases} 1 - e^{-\lambda \cdot x}, & \textit{falls } x \geq 0 \\ 0, & \textit{falls } x < 0 \end{cases} .$$

Bez. 4 *Die Zufallsvariable X heißt exponentialverteilt, bez. $X \sim Exp(\lambda)$.*

Die Dichtefunktion ist

$$f(x) = F'(x) = \begin{cases} \lambda \cdot e^{-\lambda \cdot x}, & \textit{falls } x \geq 0 \\ 0, & \textit{falls } x < 0 \end{cases}.$$

Weiterhin gilt:

$$\lim_{x \to -\infty} F(x) = 0, \qquad \lim_{x \to +\infty} F(x) = 1.$$

Bsp. 2.6 *Sei die Zufallsvariable*

$$X: \quad (\Omega, \mathcal{E}, P) \to (\mathbb{R}^1, \mathcal{B}^1, P_X)$$

der Meßfehler bei Messung einer physikalischen Konstanten.

Der W.raum $(\Omega, \mathcal{E}, P)$ *ist ein Modell eines im Hintergrund wirkenden Zufallsmechanismus, der nicht näher beschrieben werden kann,*
Fehler im Meßinstrument
zufällige äußere Einflüsse.

Er enthält alle nicht näher bestimmbaren zufälligen Effekte. Zur Beschreibung dient der Bildraum $(\mathbb{R}^1, \mathcal{B}^1, P_X)$.

Oft kann man annehmen,

$$P_X(B) = \frac{1}{\sqrt{2\pi}\sigma} \int_B e^{-\frac{1}{2}\left(\frac{t-\mu}{\sigma}\right)^2} dt.$$

Die Zufallsvariable X mit der Verteilungsfunktion

$$F(x) = \frac{1}{\sqrt{2\pi}\sigma} \int\limits_{-\infty}^{x} e^{-\frac{1}{2}\left(\frac{t-\mu}{\sigma}\right)^2} dt.$$

heißt normalverteilt mit den Parametern (μ, σ^2), bez. $X \sim \mathcal{N}(\mu, \sigma^2)$. Die zugehörige Dichtefunktion hat die Form:

$$f(x) = \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2}, \quad \sigma > 0.$$

Ist $f(x)$ wirklich eine Dichtefunktion? Offensichtlich ist $f(x) \geq 0$ für alle Zahlen $x \in \mathbb{R}$ und $\sigma > 0$. Es bleibt zu untersuchen, ob gilt:

$$\lim_{x\to\infty} F(x) = \int\limits_{-\infty}^{+\infty} \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{1}{2}\left(\frac{t-\mu}{\sigma}\right)^2} dt = \int\limits_{-\infty}^{+\infty} f(t)\, dt = 1.$$

Wir bezeichnen

$$\int\limits_{-\infty}^{+\infty} \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2} dx =: I.$$

Wir betrachten zunächst:

$$
\begin{aligned}
I^2 &= \left(\frac{1}{\sqrt{2\pi}\sigma} \int\limits_{-\infty}^{+\infty} e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2} \, dx \right)^2 \\
&= \frac{1}{2\pi\sigma^2} \left(\int\limits_{-\infty}^{+\infty} e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2} \, dx \right) \left(\int\limits_{-\infty}^{+\infty} e^{-\frac{1}{2}\left(\frac{y-\mu}{\sigma}\right)^2} \, dy \right) \\
&= \frac{1}{2\pi\sigma^2} \int\limits_{-\infty}^{+\infty} \left(\int\limits_{-\infty}^{+\infty} e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2} \, dx \right) e^{-\frac{1}{2}\left(\frac{y-\mu}{\sigma}\right)^2} \, dy \\
&= \frac{1}{2\pi\sigma^2} \int\limits_{-\infty}^{+\infty} \int\limits_{-\infty}^{+\infty} e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2} e^{-\frac{1}{2}\left(\frac{y-\mu}{\sigma}\right)^2} \, dx \, dy
\end{aligned}
$$

Wir führen nun eine Substitution durch:

$$
s := \frac{x-\mu}{\sigma} \qquad t := \frac{y-\mu}{\sigma}.
$$

Dann gilt:

$$
x = s\sigma + \mu \qquad y = t\sigma + \mu,
$$

$$
dx = \sigma \, ds \qquad dy = \sigma \, dt.
$$

Wir erhalten damit:

$$\begin{aligned} I^2 &= \frac{1}{2\pi\sigma^2} \int\limits_{-\infty}^{+\infty} \int\limits_{-\infty}^{+\infty} e^{-\frac{1}{2}s^2} e^{-\frac{1}{2}t^2} \sigma^2 \, ds \, dt \\ &= \frac{1}{2\pi} \int\limits_{-\infty}^{+\infty} \int\limits_{-\infty}^{+\infty} e^{-\frac{1}{2}(s^2+t^2)} \, ds \, dt \end{aligned}$$

Wir führen eine weitere Substitution durch, Polarkoordinaten:

$$s = r\cos\varphi \qquad t = r\sin\varphi.$$

Dann gilt allgemein nach der Substitutionsregel:

$$\int\int g(s,t)\,ds\,dt = \int\int g(r,\varphi) \det J \, dr\, d\varphi,$$

$$\det J = |J| = \begin{vmatrix} \frac{\partial s}{\partial r} & \frac{\partial s}{\partial \varphi} \\ \frac{\partial t}{\partial r} & \frac{\partial t}{\partial \varphi} \end{vmatrix} = \begin{vmatrix} \cos\varphi & -r\sin\varphi \\ \sin\varphi & r\cos\varphi \end{vmatrix}$$

$$= r\cos^2\varphi + r\sin^2\varphi$$

$$= r(\cos^2\varphi + \sin^2\varphi) = r$$

$$I^2 = \frac{1}{2\pi}\int_0^{2\pi}\int_0^{\infty} e^{-\frac{1}{2}(r^2\cos^2\varphi + r^2\sin^2\varphi)} r\,dr\,d\varphi$$

$$= \frac{1}{2\pi}\int_0^{2\pi}\int_0^{\infty} e^{-\frac{1}{2}r^2} r\,dr\,d\varphi$$

$$= \frac{1}{2\pi}\int_0^{2\pi} \left[-e^{-\frac{r^2}{2}}\right]_0^{\infty} d\varphi$$ (durch Differentiation leicht nachvollziehbar!)

$$= \frac{1}{2\pi}\int_0^{2\pi} d\varphi = \frac{1}{2\pi}2\pi = 1$$

$\Longrightarrow$ $I = 1$, d.h. f ist eine Dichtefunktion.

Zusammenfassung (Zufallsvariable).

Eine (meßbare) Abbildung

$$X : \quad \Omega \longrightarrow \mathbb{R}$$

heißt Zufallsvariable.

Jedem Element ω des Stichprobenraumes Ω wird eine reelle Zahl zugeordnet.

Def.: Die Zufallsvariable X heißt diskret, wenn X nur endlich viele oder abzählbar unendlich viele Werte x_i annehmen kann. Jeder dieser Werte kann mit einer gewissen Wkt. $p_i = P(X = x_i)$ auftreten.

Bsp.: - geografische Lage (N,O,S,W)
- Länge einer Warteschlange
- Anzahl der erreichten Punkte in der Klausur.

Def.: Die Zufallsvariable X heißt <u>stetig</u>, falls X beliebige Werte in einem Intervall (a, b), $[a, b]$, $(a, b]$, $(a, b]$, $(-\infty, a)$, (b, ∞), $(-\infty, a]$, $[b, \infty)$, $(-\infty, \infty)$ annehmen kann.

Bem.: Jeder einzelne Wert $x_i \in (a, b)$ (oder in einem der anderen Intervalle) hat die Wkt. Null.

Die <u>Verteilungsfunktion</u> F wird dann durch die sogen. <u>Dichtefunktion</u> f beschrieben,

$$F(x) = P(X \leq x) = \int_{-\infty}^{x} f(t)\,dt$$