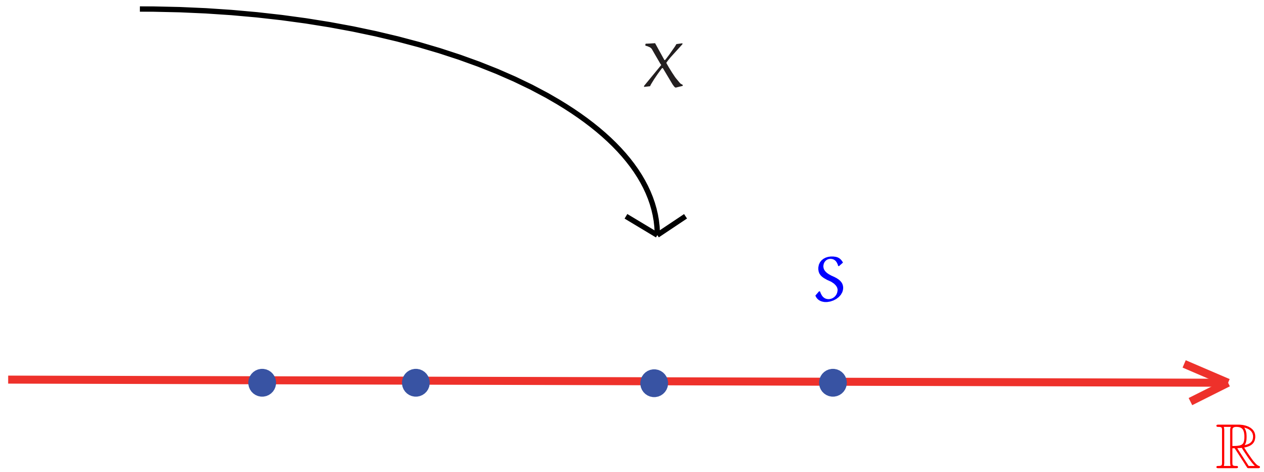
Vorlesung 3b

Der Erwartungswert

von diskreten reellwertigen Zufallsvariablen

Teil 2

X sei eine diskrete reellwertige Zufallsvariable



$$\mathbf{E}[X] := \sum_{a \in S} a \, \mathbf{P}(X = a) \, .$$

heißt *Erwartungswert* von X.

Mit X und Y ist auch deren Summe

$$X + Y$$

eine rellwertige Zufallsvariable.

Sind X und Y diskret, so auch $X + Y$.

Die wichtigste Eigenschaft des Erwartungswerts ist die

<u>Additivität</u>

$$\mathbf{E}[X + Y] = \mathbf{E}[X] + \mathbf{E}[Y]$$

Wir werden sie in Kürze herleiten;

erst erinnern wir uns an die Anwendungsbeispiele aus VL 3a:

BEISPIEL 1

Erwartungswert der Binomialverteilung

X sei Bin(n, p)-verteilt.

$$\boxed{\mathbf{E}[X] = np}$$

BEISPIEL 2

Erwartungswert der hypergeometrischen Verteilung

X sei Hyp(n, g, w)-verteilt.

$$\boxed{\mathbf{E}[X] = n\frac{w}{g}}$$

BEISPIEL 3

Runs beim Münzwurf

$Z := (Z_1, Z_2, ..., Z_n)$ n-facher p-Münzwurf

$$\mathbf{P}(Z_i = 1) = p \qquad \mathbf{P}(Z_i = 0) = q := 1 - p$$

Run: ein Block von Nullen (Einsen),

der nicht echt in einem größeren Block enthalten ist

$R :=$ Anzahl Runs in Z

$$00000000 \qquad R = 1$$

$$11100011 \qquad R = 3$$

$$10101010 \qquad R = 8$$

$$\mathbf{E}[R] = ?$$

Dazu schreiben wir R als Summe von Zählern.

Bei jedem Wurf zählen wir eins dazu,

wenn bei diesem Wurf ein Run beginnt:

$Y_i := 1$ falls bei i ein Run beginnt, $\quad Y_i := 0$ sonst

$$R = Y_1 + Y_2 + \ldots + Y_n$$

$$Y_1 \equiv 1$$

$$\{Y_i = 1\} = \{(Z_{i-1}, Z_i) = (0,1) \text{ oder } (1,0)\} \qquad (i > 1)$$

$$\mathbf{P}(Y_i = 1) = qp + pq \qquad (i > 1)$$

$$\mathbf{E}[Y_i] = 2pq \qquad (i > 1)$$

$$\mathbf{E}[R] = \mathbf{E}[Y_1] + \mathbf{E}[Y_2] + \mathbf{E}[Y_3] + \ldots + \mathbf{E}[Y_n]$$

$$\boxed{\mathbf{E}[R] = 1 + 2pq(n-1)}$$

Gerade haben wir die Linearität des Erwartungswertes

schon in Beispielen angewendet.

Wir werden sie jetzt aus der Definition

$$\mathbf{E}[X] := \sum_{a \in S} a\, \mathbf{P}(X = a)$$

herleiten.

Hilfreich dabei ist

eine wichtige „Transformationsformel“:

(Buch S. 23)

Sei X diskrete Zufallsvariable mit $\mathbf{P}(X \in S) = 1$

und h eine Abbildung von S nach $\mathbb{R}$

(so dass der Erwartungswert der Zufallsvariablen $h(X)$

wohldefiniert ist). Dann ist

$$\mathbf{E}[h(X)] = \sum_{a \in S} h(a)\,\mathbf{P}(X = a)\ .$$

Beweis.

$$\sum_{b \in h(S)} b \, \mathbf{P}(h(X) = b)$$

$$= \sum_{b \in h(S)} b \sum_{a \in h^{-1}(b)} \mathbf{P}(X = a)$$

$$= \sum_{b \in h(S)} \sum_{a \in h^{-1}(b)} h(a) \, \mathbf{P}(X = a)$$

$$= \sum_{a \in S} h(a) \, \mathbf{P}(X = a) \, . \quad \square$$

Satz [**Linearität des Erwartungswertes**]

(Buch S. 52)

Für reellwertige Zufallsvariable X_1, X_2

mit wohldefiniertem Erwartungswert gilt

$$\mathbf{E}[c_1 X_1 + c_2 X_2] = c_1 \mathbf{E}[X_1] + c_2 \mathbf{E}[X_2]\,, \quad c_1, c_2 \in \mathbb{R}\,.$$

Beweis.

Seien $S_1, S_2 \subset \mathbb{R}$ abzählbar mit

$$\mathbf{P}(X_1 \in S_1) = \mathbf{P}(X_2 \in S_2) = 1.$$

Aus der Transformationsformel folgt mit

$$h(a_1, a_2) := c_1 a_1 + c_2 a_2:$$

$$\mathbf{E}[c_1X_1 + c_2X_2]$$

$$= \sum_{a_1 \in S_1} \sum_{a_2 \in S_2} (c_1a_1 + c_2a_2)\, \mathbf{P}(X_1 = a_1, X_2 = a_2)$$

$$= c_1 \sum_{a_1 \in S_1} a_1 \sum_{a_2 \in S_2} \mathbf{P}(X_1 = a_1, X_2 = a_2)$$

$$+ c_2 \sum_{a_2 \in S_2} a_2 \sum_{a_1 \in S_1} \mathbf{P}(X_1 = a_1, X_2 = a_2)$$

$$\sum_{a_1 \in S_1} a_1 \sum_{a_2 \in S_2} \mathbf{P}(X_1 = a_1, X_2 = a_2)$$

$$= \sum_{a_1 \in S_1} a_1 \mathbf{P}(X_1 = a_1)$$

$$= \mathbf{E}[X_1] \qquad \square$$

Zusammenfassung

1.

Was ist der Erwartungswert?

$$\mathbf{E}[X] = \sum x\mathbf{P}(X = x)$$

und

$$\mathbf{E}[X] = \lim \frac{X_1 + ... + X_n}{n}$$

für “unabhängige Wiederholungen” $X_1, X_2, \ldots$

2.

Was ist die wichtigste Eigenschaft des Erwartungswertes?

Die Linearität:

$$\mathbf{E}[\alpha X + \beta Y] = \alpha \mathbf{E}[X] + \beta \mathbf{E}[Y]$$

3.

Wie berechnet man $\mathbf{E}[X]$ am besten?

Oft dadurch,

dass man X als Summe schreibt:

$$X = Z_1 + ... + Z_n$$

$$\mathbf{E}[X] = \mathbf{E}[Z_1] + ... + \mathbf{E}[Z_n]$$

Das Rechnen mit

Indikatorvariablen und Ereignissen,

Erwartungswerten und

Wahrscheinlichkeiten.

Wir wissen schon:

Eriegnisse kann man oft

auf verschiedene Weise darstellen:

Beispiel 1:

Sei (X_1, X_2) das Paar der Augenzahlen

beim zweimaligen (gewöhnlichen) Würfeln.

$\{X_1 = 3\}$

$= \{(X_1, X_2) \in \{(3,1), (3,2), (3,3), (3,4), (3,5), (3,6)\}\}$

$\{X_1 + X_2 = 3\} = \{(X_1, X_2) \in \{(1,2), (2,1)\}\}$

Beispiel 2: Kollisionen (vgl. Vorlesung 1b):

Wir stellen uns vor, dass die Individuen

der Reihe nach ihr Kennzeichen bekommen.

X_i ... zufälliges Kennzeichen des i-ten Individuums

T sei der Moment der ersten Kollision:

$$T = \min\{i \geq 1 : X_i = X_j \text{ für ein } j < i\}\,.$$

Dann gilt für das Ereignis

"keine Kollision unter den ersten n Individuen":

$$\{X_i \neq X_j \text{ für } j < i \leq n\} = \{T > n\}\,.$$

Allgemeiner gilt

für die “Verarbeitung” $h(X)$ einer Zufallsvariablen X:

$$\{h(X) \in B\} = \{X \in h^{-1}(B)\}.$$

Wir erinnern an eine Definition aus Vorlesung 3b

(vgl. Buch S. 36-38):

Für eine Zufallsvariable X mit Wertebereich S und $A \subset S$ betrachten wir die Zufallsvariable $\mathbf{1}_A(X)$.

Sie nimmt den Wert 1 an, wenn das Ereignis $\{X \in A\}$ eintritt, und den Wert 0, wenn das Ereignis $\{X \in A^c\}$ eintritt.

$$I_{\{X \in A\}} := \mathbf{1}_A(X)$$

heißt *Indikatorvariable* des Ereignisses $\{X \in A\}$.

$$\{X \in A\} = \{I_{\{X \in A\}} = 1\},$$

Für jedes Ereignis E gilt: $E = \{I_E = 1\}$.

Ereignisse sind gleich,

wenn das für ihre Indikatorvariablen zutrifft.

Und mit Indikatorvariablen lässt sich leicht rechnen.

Für jede S-wertige Zufallsvariable X gilt:

$$I_{\{X \in S\}} = \mathbf{1}_S(X)$$

ist eine Zufallsvariable, die “sicher” auf den Ausgang 1 fällt.

$$I_{\{X \in \emptyset\}} = \mathbf{1}_\emptyset(X)$$

ist eine Zufallsvariable, die “sicher” auf den Ausgang 0 fällt.

Das *sichere Ereignis* E_s ist dadurch charakterisiert, dass seine Indikatorvariable sicher den Wert 1 annimmt.

$$I_{E_s} = 1$$

Das *unmögliche Ereignis* E_u ist dadurch charakterisiert, dass seine Indikatorvariable sicher den Wert 0 annimmt:

$$I_{E_u} = 0.$$

Für zwei Ereignisse E_1, E_2

hat deren "Oder-Ereignis" die Indikatorvariable

$$I_{E_1 \cup E_2} := \max(I_{E_1}, I_{E_2})$$

und deren "Und-Ereignis" die Indikatorvariable

$$I_{E_1 \cap E_2} := \min(I_{E_1}, I_{E_2}).$$

Man spricht auch von "Vereinigung" und "Durchschnitt"

der Ereignisse E_1 und E_2.

Für $b_1, b_2 \in \{0, 1\}$ gilt die Identität:

$$b_1 + b_2 = \max(b_1, b_2) + \min(b_1, b_2).$$

Dies überträgt sich auf

$$I_{E_1} + I_{E_2} = I_{E_1 \cup E_2} + I_{E_1 \cap E_2}$$

Falls

$$E_1 \cap E_2 = E_u ,$$

so heißen E_1 und E_2

disjunkte oder *sich ausschließende Ereignisse*.

Gilt $E_1 = E_1 \cap E_2$, so schreiben wir

$$E_1 \subset E_2 \, .$$

und sagen:

“Mit E_1 tritt sicher auch E_2 ein”

oder auch

“Das Ereignis E_1 zieht das Ereignis E_2 nach sich.”

Für jedes Ereignis E ist sein *Komplementärereignis*

$$E^c$$

definiert durch

$$I_{E^c} := 1 - I_E \quad \text{bzw.} \quad E^c := \{I_E = 0\}\,.$$

Wegen $\mathbf{1}_{A^c} = 1 - \mathbf{1}_A$ gilt

$$\{X \in A\}^c = \{X \in A^c\}\,.$$

Seien X, Y Zufallsvariable mit demselben Wertebereich S.

$D := \{(x, y) \in S^2 : x = y\}$, die „Diagonale“ in S^2.

$$\{X = Y\} := \{(X, Y) \in D\}$$

bzw.

$$I_{\{X=Y\}} = \mathbf{1}_D(X, Y)$$

Für reellwertige Zufallsvariable X, Y setzen wir

$$\{X \leq Y\} := \{(X, Y) \in H\}$$

mit $H := \{(x, y) \in \mathbb{R}^2 : x \leq y\}$,

der Halbraum über (und einschließlich) der Diagonalen.

Wir schreiben

$$X \leq Y \quad :\Leftrightarrow \quad \{X \leq Y\} = E_S \; .$$

Bringen wir jetzt wieder Wahrscheinlichkeiten ins Spiel.

Für Indikatorvariablen und Ereignisse

gilt die Beziehung

$$\mathbf{E}[Z] = \mathbf{P}(Z = 1),$$

$$\mathbf{E}[I_E] = \mathbf{P}(I_E = 1) = \mathbf{P}(E).$$

Aus dem Rechnen mit Indikatorvariablen

und aus der Linearität des Erwartungswertes

ergeben sich die Regeln

für das Rechnen mit Wahrscheinlichkeiten

(Buch S. 57-58):

(i) $\mathbf{P}(E_s) = 1, \qquad \mathbf{P}(E_u) = 0.$

(ii) $\mathbf{P}(E_1) + \mathbf{P}(E_2) = \mathbf{P}(E_1 \cup E_2) + \mathbf{P}(E_1 \cap E_2),$

insbesondere $\mathbf{P}(E_1 \cup E_2) \leq \mathbf{P}(E_1) + \mathbf{P}(E_2).$

(iii) $\mathbf{P}(E_1 \cup E_2) = \mathbf{P}(E_1) + \mathbf{P}(E_2)$

falls E_1 und E_2 disjunkt.

(iv) $\mathbf{P}(E^c) = 1 - \mathbf{P}(E).$

(v) $\mathbf{P}(E_1) \leq \mathbf{P}(E_2)$, falls $E_1 \subset E_2$.

Bei (ii) verwendet man die Identität

$$I_{E_1} + I_{E_2} = I_{E_1 \cup E_2} + I_{E_1 \cap E_2}$$

(ii) wird verallgemeinert durch die

Einschluss-Ausschluss-Formel:

$$\mathbf{P}(E_1 \cup \cdots \cup E_n) \\ = \sum_i \mathbf{P}(E_i) - \sum_{i<j} \mathbf{P}(E_i \cap E_j) + \cdots \pm \mathbf{P}(E_1 \cap \cdots \cap E_n) \, .$$

Beweis:

$$1 - I_{E_1 \cup \cdots \cup E_n}$$

fällt genau dann als 0 aus,

wenn mindestens eines der I_{E_i} als 1 ausfällt,

ist also gleich dem Produkt

$$(1 - I_{E_1}) \cdots (1 - I_{E_n})$$

Ausmultiplizieren ergibt

$$1 - \sum_i I_{E_i} + \sum_{i<j} I_{E_i \cap E_j} - \cdots .$$

Gehe dann über zum Erwartungswert. □