Script **generated by TTT**

Title: Meixner: Zue_DWT (08.05.2014)

Date: Thu May 08 14:16:05 CEST 2014

Duration: 53:49 min

Pages: 28

SS 2014

Zentralübung zur Vorlesung Diskrete Wahrscheinlichkeitstheorie

Dr. Werner Meixner

Fakultät für Informatik
TU München

http://www14.in.tum.de/lehre/2014SS/dwt/uebung/

8. Mai 2014

Übersicht:

1. Übungsbetrieb	Fragen Zusatzaufgabe 1
2. Thema	Verteilungen
3. Vorbereitung	VA Blatt 4

Neue Korrekturregeln und Zusatzaufgabe 1

Die erhaltenen Punkte bei 4 Aufgaben A_1, A_2, A_3 bzw. A_4 eines Übungsblattes seien bei vollständiger Korrektur a_1, a_2, a_3 bzw. a_4. Die erhaltene Punktesumme ist dann $a_1 + a_2 + a_3 + a_4$.

Zur Reduzierung der Korrekturarbeit werden Laplace-zufällig 2 Aufgaben ausgewählt, diese beiden Aufgaben vollständig korrigiert und anschließend die Punktesumme P der beiden Aufgaben verdoppelt gutgeschrieben.

Berechnen Sie den Erwartungswert $\mathbb{E}[P]$ der gutgeschriebenen Punktesumme!

Lösung

Es gilt $\mathbb{E}[P] = a_1 + a_2 + a_3 + a_4$.

2. Thema: Verteilungen

Ziel: Den Zusammenhang unter gewissen Verteilungen herstellen.

2.1 Welche Verteilungen betrachten wir?

Diskrete Verteilungen, die eng mit Bernoulli-verteilten Zufallsvariablen zusammenhängen:

- Geometrische Verteilung Geo(p): $f_{Z_1}(z) = pq^{z-1}$.
- Negativ-Binomial-Verteilung NegativBin(x, p): $f_{Z_2}(z) = \binom{z-1}{x-1} p^x q^{z-x}$
- Binomialverteilung Bin(z, p): $f_{X_1}(x) = \binom{z}{x} p^x q^{z-x}$.
- Poisson-Verteilung Po(λ): $f_{X_2}(x) = \frac{e^{-\lambda}\lambda^x}{x!}$.

Dabei gilt für die Argumente der Dichtefunktionen entsprechend $z \in \mathbb{N}$ bzw. $z \in \mathbb{N}$ bzw. $x \in \mathbb{N}_0$ bzw. $x \in \mathbb{N}_0$.

Für alle übrigen Argumente aus $\mathbb{R}$ werden die Dichten gleich 0 gesetzt.

Diskrete Verteilungen, die eng mit Bernoulli-verteilten Zufallsvariablen zusammenhängen:

- Geometrische Verteilung Geo(p): $f_{Z_1}(z) = pq^{z-1}$.
- Negativ-Binomial-Verteilung NegativBin(x, p): $f_{Z_2}(z) = \binom{z-1}{x-1} p^x q^{z-x}$
- Binomialverteilung Bin(z, p): $f_{X_1}(x) = \binom{z}{x} p^x q^{z-x}$.
- Poisson-Verteilung Po(λ): $f_{X_2}(x) = \frac{e^{-\lambda}\lambda^x}{x!}$.

Dabei gilt für die Argumente der Dichtefunktionen entsprechend $z \in \mathbb{N}$ bzw. $z \in \mathbb{N}$ bzw. $x \in \mathbb{N}_0$ bzw. $x \in \mathbb{N}_0$.

Für alle übrigen Argumente aus $\mathbb{R}$ werden die Dichten gleich 0 gesetzt.

2.2 Beispiel einer Binomialverteilung

VA 4

Wir wählen nacheinander (gleichverteilt) zufällig und unabhängig Buchstaben aus der Multimenge der Buchstaben des Wortes `CHOOSE` aus.

Berechnen Sie Erwartungswert und Varianz der folgenden Zufallsvariablen Z:

Z := Anzahl der Züge (mit Zurücklegen)
bis zum dritten Mal `O` gezogen wurde.

Vergleichen Sie VA 3 von Blatt 3!

Bestimmung der Verteilung für Z

Die Aufgabe ist ein Spezialfall der Bestimmung der Wahrscheinlichkeit, dass bei der Wiederholung z der Auswertung einer Bernoulli mit Erfolgswahrscheinlichkeit p verteilten Zufallsvariablen I zum x-ten Mal das Ereignis $I = 1$ eintritt.

Die Wahrscheinlichkeitsdichte von Z ist

$$f_Z(z) = \binom{z-1}{x-1} p^x q^{z-x} .$$

Für $p = \frac{1}{3}$ und $k = 3$ folgt also

$$f_Z(z) = \binom{z-1}{2} \left(\frac{1}{3}\right)^2 \left(\frac{2}{3}\right)^{z-2} .$$

Bestimmung der Verteilung für Z

Die Aufgabe ist ein Spezialfall der Bestimmung der Wahrscheinlichkeit, dass bei der Wiederholung z der Auswertung einer Bernoulli mit Erfolgswahrscheinlichkeit p verteilten Zufallsvariablen I zum x-ten Mal das Ereignis $I = 1$ eintritt.

Die Wahrscheinlichkeitsdichte von Z ist

$$f_Z(z) = \binom{z-1}{x-1} p^x q^{z-x} .$$

Für $p = \frac{1}{3}$ und $k = 3$ folgt also

$$f_Z(z) = \binom{z-1}{2} \left(\frac{1}{3}\right)^2 \left(\frac{2}{3}\right)^{z-2} .$$

Berechnung $\mathbb{E}[Z]$ und $\mathrm{Var}[Z]$

Wir berechnen nicht direkt, sondern nutzen die folgende Darstellung:

Es gilt

$$Z = Z_1 + Z_2 + Z_2 ,$$

wobei Z_1, Z_2, Z_3 unabhängige und geometrisch verteilte Zufallsvariable sind mit $p = \frac{1}{3}$.

Dann gilt $\mathbb{E}[Z_i] = 3$ und $\mathrm{Var}[Z_i] = \frac{q}{p^2} = \frac{2}{3} \cdot 9 = 6$.

Ergebnis:

$$\mathbb{E}[Z] = 3 \cdot 3 = 9 \quad \text{und} \quad \mathrm{Var}[Z] = 3 \cdot 9 = 27 .$$

2.3 Beispiel einer geometrischen Verteilung

TA 4 von Blatt 3

2.4 Das Konzept der Wiederholung bei Zufallsvariablen

Viele wahrscheinlichkeitstheoretische Experimente werden durch unabhängige Wiederholung eines bestimmten Experiments definiert.

Dem entspricht die Betrachtung einer (unendlichen) Folge von unabhängigen Zufallsvariablen $I_{p,1}, I_{p,2}, \ldots, I_{p,n}, \ldots$, die wie folgt definiert werden:

Definition:

Sei I_p eine Bernoulli-verteilte Zufallsvariable über einem Wahrscheinlichkeitsraum $(\Omega, \Pr)$ mit Erfolgswahrscheinlichkeit p.

Für alle $n \in \mathbb{N}$ wird durch die unabhängige n-te Wiederholung der Auswertung von I_p eine Zufallsvariable $I_{p,n}$ definiert.

„Aussage":

Insgesamt erhält man ein unabhängiges System von unendlich vielen Zufallsvariablen mit gleicher Verteilung wie I_p

$$I_{p,1}, I_{p,2}, \ldots, I_{p,n}, \ldots \quad .$$

Kritik der Definition bzw. „Aussage" zur Unabhängigkeit:

1. Systeme von „unabhängigen" Zufallsvariablen setzen einen gemeinsamen Wahrscheinlichkeitsraum voraus.

   Aber!

   Auf $(\Omega, \Pr)$ sind alle $I_{p,i}$ identisch und insbesondere abhängig.

   Dies kann also nicht! der geforderte gemeinsame Wahrscheinlichkeitsraum sein.

Was ist zu tun?

1. Lösungsmöglichkeit:

Man betrachtet stets nur endliche Abschnitte der Folge und kann dann für diese Folge den W'keitsraum als endliches Produkt der Räume für die beteiligten Bernoulli-verteilten Variablen.
In diesem Fall kommt man mit der bisherigen Definition der W'keitsräume aus.

2. Lösungsmöglichkeit:

Wir definieren ein neues Konzept für Wahrscheinlichkeitsräume nach Kolmogorov, in dem es keine atomaren Ereignisse gibt.

2.5 Gemeinsame Herleitung der Verteilungen

Wir betrachten die unbegrenzte (∞-fache) Wiederholung eines Experiments mit den Ereignissen eines Wahrscheinlichkeitsraumes $(\Omega, \Pr)$ und der Bewertung der Ereignisse durch I_p

wie folgt:

Bei unbegrenzter Wiederholung des Experiments erhalten wir eine Folge von Ergebnissen $\omega_1, \omega_2, \ldots$ mit Bewertungen $I_p(\omega_1), I_p(\omega_2), \ldots$.

Dann können wir eine unendliche Folge von Zufallsvariablen $I_{p,1}, I_{p,2}, \ldots$ definieren durch Abbildungen $I_{p,i} : \Omega^{\mathbb{N}} \to \{0,1\}$ mit

$$I_{p,i}((\omega_1, \omega_2, \ldots)) = I_p(\omega_i)\,.$$

Sei $Y = I_{p,1}, I_{p,2}, \ \ldots\ , I_{p,z}, \ \ldots$.

Sei $A_{x,z}$ die Menge aller Folgen Y aus $\Omega^{\mathbb{N}}$ mit der Eigenschaft, dass in der Folge Y
an z-ter Stelle das x-te Mal eine 1 aufgetreten ist.

In der Sprache von Wahrscheinlichkeitsräumen würden wir sagen: $A_{x,z}$ ist das Ereignis über $\Omega^{\mathbb{N}}$, dass in der Folge Y
an z-ter Stelle das x-te Mal eine 1 aufgetreten ist.

In jedem sinnvoll zu Grunde gelegten W'keitsraum über endlichen Abschnitten der Folgen (sowie auch in später eingeführten W'keitsräumen nach Kolmogorov) gilt dann

$$Pr[A_{x,z}] = \binom{z-1}{x-1} p^x q^{z-x}\,.$$

Bemerkung: Die Mengen (Ereignisse) $A_{x,z}$ sind disjunkt.

Matrix der negativ binomialverteilten Wahrscheinlichkeiten $Pr[A_{x,z}]$:

$x =$ \ $z =$	1	2	3	4	…	k
1	p	$\binom{1}{0}pq$	$\binom{2}{0}pq^2$	$\binom{3}{0}pq^3$	…	
2	0	p^2	$\binom{2}{1}p^2q$	$\binom{3}{1}p^2q^2$	…	
3	0	0	p^3	$\binom{3}{2}p^3q$	…	
4	0	0	0	p^4	…	
⋮					⋮	
i					…	$\binom{k-1}{i-1}p^iq^{k-i}$

Spaltensumme:

Für alle $k \geq 1$ gilt

$$\sum_{i=1}^{\infty} \binom{k-1}{i-1} p^i q^{k-i} \; = \; p\,.$$

Wenn wir X_k definieren als Anzahl der Einsen im Vektor $(I_{p,1}, I_{p,2}, \ldots, I_{p,k-1})$ unter der Bedingung, dass $I_{p,k} = 1$ gilt, dann ist $X_k - 1$ binomialverteilt.

Zeilensumme:

Für alle $i \geq 1$ gilt

$$\sum_{k=1}^{\infty} \binom{k-1}{i-1} p^i q^{k-i} = 1 .$$

Wenn wir Z_i definieren als das minimale k, so dass der Vektor $(I_{p,1}, I_{p,2}, \ldots, I_{p,k})$ genau i Einsen enthält, dann ist Z_i negativ binomialverteilt.

Für $i = 1$ ergibt sich die geometrische Verteilung und

$$\sum_{k=1}^{\infty} p q^{k-1} = 1 .$$

Aus der Folge $Y = I_{p,1}, I_{p,2}, \ \ldots \ , I_{p,z}, \ \ldots$ kann man bei festgehaltener Zeile i auch i geometrisch verteilte Zufallsvariable X_j definieren. Dabei sei X_j die Anzahl der Schritte ausgehend von der $(j-1)$ten Eins bis zum Auftreten der $j$ten Eins in der Folge Y.

Dann gilt $Z_i = \sum_{j=1}^{i} X_j$.

Beachte: Z_i ist also eine Summe von geometrisch verteilten Zufallsvariablen X_j, mit der Konsequenz, dass man den Erwartungswert von Z_i als Summe von Erwartungswerten der geometrisch verteilten Zufallsvariablen X_j berechnen kann.

3. Vorbereitung VA Blatt 4

3.1 VA 1

In der statistischen Physik pflegt man die Verteilung gewisser Teilchen (Moleküle, Photonen, Elektronen, usw.) zu betrachten. Man nimmt an, dass sich jedes Teilchen durch einen Vektor von Kenngrößen charakterisieren läßt, der der Zuordnung des Teilchens in eine von N ,,Zellen" eines ,,Phasenraumes" entspricht. Der Zustand eines physikalischen Systems wird dann dadurch beschrieben, dass man angibt, wie viele Teilchen sich jeweils in einer Zelle befinden.

Die Annahme, dass jede Lage eines n Gasmoleküls in N Zellen gleichwahrscheinlich ist, ist Basis der *Maxwell-Boltzmannsche Statistik*. Dabei wird die Unterscheidbarkeit der Gasmoleküle vorausgesetzt.

3. Vorbereitung VA Blatt 4

3.1 VA 1

In der statistischen Physik pflegt man die Verteilung gewisser Teilchen (Moleküle, Photonen, Elektronen, usw.) zu betrachten. Man nimmt an, dass sich jedes Teilchen durch einen Vektor von Kenngrößen charakterisieren läßt, der der Zuordnung des Teilchens in eine von N ,,Zellen" eines ,,Phasenraumes" entspricht. Der Zustand eines physikalischen Systems wird dann dadurch beschrieben, dass man angibt, wie viele Teilchen sich jeweils in einer Zelle befinden.

Die Annahme, dass jede Lage eines n Gasmoleküls in N Zellen gleichwahrscheinlich ist, ist Basis der *Maxwell-Boltzmannsche Statistik*. Dabei wird die Unterscheidbarkeit der Gasmoleküle vorausgesetzt.

Bestimmen Sie die Wahrscheinlichkeit $p(k,n)$, dass sich genau k Gasmoleküle in einer beliebigen Zelle Z der N Zellen befinden, einer Binomialverteilung genügen und zeigen Sie dazu

$$p(k,n) = b(k; n, \tfrac{1}{N}) := \binom{n}{k} \cdot \left(\frac{1}{N}\right)^k \cdot \left(\frac{N-1}{N}\right)^{n-k}.$$

Lösung

Jede örtliche Zuordnung der n Gasmoleküle $G_1, \ldots, G_n$, so dass sich jedes Molekül in einer der N Zellen $Z_1, \ldots, Z_N$ befindet, stellt ein Elementarereignis dar.

Wir codieren die Gasmoleküle durch natürliche Zahlen $1, 2, \ldots, n$ und die Zellen durch natürliche Zahlen $1, 2, \ldots, N$. Die Elementarereignisse $e \in \Omega$ codieren wir durch n-Tupel $e = (e_1, e_2, \ldots, e_n)$ mit $e_i \in [N]$ und der Bedeutung, dass der Wert der i-ten Komponente von e die Zelle liefert, in der sich das Molekül G_i befindet.

Es gilt mit Anwendung des Laplace Prinzips

$$e \in \Omega = \{(e_1, e_2, \ldots, e_n); e_i \in [N]\}, \ \Pr[e] = N^{-n}.$$

Wir betrachten nun eine bestimmte Zelle Z.

Für jedes der n Moleküle definieren wir eine Bernoulli Variable X_i, die angibt, ob sich bei Ereignis e das Molekül G_i in der Zelle Z befindet oder nicht. Die Zufallsvariablen X_i sind Bernoulli-verteilt mit Erfolgswahrscheinlichkeit $p = \frac{1}{N}$.
Die Zufallsvariablen $X_1, X_2, \ldots, X_n$ sind unabhängig.

Nun betrachten wir die Zufallsvariable

$$X_Z = \sum_{i=1}^{n} X_i .$$

X_Z ist binomialverteilt, i. Z. $X \sim \mathrm{Bin}(n, \frac{1}{N})$, mit Dichtefunktion

$$f_{X_Z}(k) = b(k; n, \tfrac{1}{N}) := \binom{n}{k} \cdot \left(\frac{1}{N}\right)^k \cdot \left(\frac{N-1}{N}\right)^{n-k}.$$

$f_{X_Z}(k)$ ist die gesuchte Wahrscheinlichkeit.

Nun betrachten wir die Zufallsvariable

$$X_Z = \sum_{i=1}^{n} X_i \,.$$

X_Z ist binomialverteilt, i. Z. $X \sim \mathrm{Bin}(n, \frac{1}{N})$, mit Dichtefunktion

$$f_{X_Z}(k) = b(k; n, \tfrac{1}{N}) := \binom{n}{k} \cdot \left(\frac{1}{N}\right)^k \cdot \left(\frac{N-1}{N}\right)^{n-k} \,.$$

$f_{X_Z}(k)$ ist die gesuchte Wahrscheinlichkeit.