4. diskrete Verteilungen

Diskrete Stochastik

Prof. Dr. Andreas Vogt

n|w Fachhochschule Nordwestschweiz

diskrete Verteilungen

Wir schauen uns nun die wichtigsten diskreten Verteilungen an.

Diskret bedeutet, dass die Zufallsvariablen nur Werte aus einer endlichen oder abzählbaren Menge annehmen.

Etwa:
- Ergebnisse beim Würfeln
- Anzahl Münzwürfe bis zum ersten Mal Kopf

$\vdots$

Nicht diskret verteilt ist etwa die Lebensdauer eines Bauteils.

Binomial-Verteilung

Definition:

Die Verteilung einer Zufallsvariable X, die die Anzahl an Treffern bei der n-maligen unabhängigen Durchführung eines Experiments mit zwei Ausgängen, *Treffer* und *kein Treffer*, wobei die p die Wahrscheinlichkeit für *Treffer* bezeichnet, heisst Binomial-verteilt mit Parametern n, p.

Matlab-Funktionen:

Dichte: `binopdf`(k, n, p)

Verteilungsfunktion: `binocdf`(k, n, p)

Schreibweise: $X \sim Bin(n, p)$

Dichte von X: $f(k) = \binom{n}{k} p^k (1-p)^{n-k}$, $k = 0, 1, \ldots, n$

$E(X) = n \cdot p$

$V(X) = n \cdot p \cdot (1-p)$

Begründung für die Dichte:

Modellierung als n unabhängige Versuche mit $\Omega_i = \{$*Treffer*, *kein Treffer*$\}$ und $f_i($*Treffer*$) = p$ und $f_i($*kein Treffer*$) = 1 - p$, also $\Omega_{\text{gesamt}} = \{$*Treffer*, *kein Treffer*$\}^n$, $f(\omega_1, \ldots, \omega_n) = f_1(\omega_1) \cdot \ldots \cdot f_n(\omega_n)$.

Im Fall $n = 3$ gilt etwa: $f($*Treffer*, *kein Treffer*, *Treffer*$) = p \cdot (1-p) \cdot p$

Oder: $f($*kein Treffer*, *Treffer*, *Treffer*$) = (1-p) \cdot p \cdot p$

Wenn also in n Versuchen k Treffer auftreten, also $n - k$ Nicht-Treffer, dann erhalten wir dafür die Wahrscheinlichkeit: $\binom{n}{k} p^k (1-p)^{n-k}$

Bernoulli-Verteilung

Definition:

Die Binomialverteilung mit $n = 1$ und Parameter p heisst Bernoulli-verteilt mit Parameter p.

Schreibweise: $X \sim B(p)$

Dichte von X: $f(0) = 1 - p$, $f(1) = p$.

$E(X) = p$

$$E(X) = 0 \cdot (1 - p) + 1 \cdot p = p$$

$V(X) = p(1 - p)$

$$V(X) = (0 - p)^2 \cdot (1 - p) + (1 - p)^2 \cdot p = p(1 - p)(p + (1 - p)) = p(1 - p)$$

Binomial-Verteilung

Beispiel:

Ein Multiple Choice Test besteht aus 12 Fragen mit je 4 Antworten und jeweils genau einer richtigen Antwortmöglichkeit. Der Test wird durch Erraten ausgefüllt. Wie gross ist die Wahrscheinlichkeit für mehr als 8 richtige Antworten?

X zähle die Anzahl der Treffer in den 12 unabhängigen Versuchen.

Dann ist $X \sim Bin\left(12, \frac{1}{4}\right)$, also

$$P(X \in \{9, 10, 11, 12\}) = P(X = 9) + \ldots + P(X = 12) = \binom{12}{9}\left(\frac{1}{4}\right)^9\left(\frac{3}{4}\right)^3 + \ldots + \binom{12}{12}\left(\frac{1}{4}\right)^{12}\left(\frac{3}{4}\right)^0$$
$$= 1 - binocdf(8, 12, 1/4) \approx 0.04\%$$

geometrische Verteilung

Definition:

Die Verteilung einer Zufallsvariable X, die die Anzahl der Versuche bis zum ersten Treffer bei der wiederholten unabhängigen Durchführung eines Experiments mit zwei Ausgängen, *Treffer* und *kein Treffer*, wobei p die Wahrscheinlichkeit für *Treffer* bezeichnet, heisst geometrisch verteilt mit Parameter p.

Schreibweise: $X \sim Geo(p)$

Dichte von X: $f(k) = (1-p)^{k-1}p$, $k = 1, 2, 3, \ldots$

$X = k$ bedeutet, dass die ersten $k-1$ Versuche jeweils kein Treffer war, und der k-te Versuch ein Treffer war.

$E(X) = \frac{1}{p}$

$V(X) = \frac{1-p}{p^2}$

Matlab-Funktionen:

Dichte: `geopdf`$(k-1, p)$

Verteilungsfunktion: `geocdf`$(k-1, p)$

Beispiel:

Wir würfeln solange, bis eine sechs kommt. Wie hoch ist die Wahrscheinlichkeit, dass dies im zehnten Versuch passiert?

Für die Zufallsvariable X, die den ersten Versuch mit Treffer angibt, gilt $X \sim Geo\left(\frac{1}{6}\right)$.

Es gilt $P(X = 10) = \left(\frac{5}{6}\right)^9 \cdot \frac{1}{6} = geopdf(9, \frac{1}{6}) \approx 3.2\%$.

hypergeometrische Verteilung

Definition:

Die Verteilung einer Zufallsvariable X, die beim n-maligen Ziehen ohne Zurücklegen aus einer Urne mit N Kugeln, von denen M eine bestimmte Eigenschaft haben, die Anzahl der Kugeln mit dieser Eigenschaft zählt, heisst hypergeometrisch verteilt mit Parametern n, N und M.

Matlab-Funktionen:
Dichte: $\texttt{hygepdf}(k, N, M, n)$
Verteilungsfunktion: $\texttt{hygecdf}(k, N, M, n)$

Schreibweise: $X \sim Hyp(N, M, n)$

Dichte von X: $f(k) = \frac{\binom{M}{k}\binom{N-M}{n-k}}{\binom{N}{n}}, \; k = 0, 1, 2, \ldots, n$

$X = k$ bedeutet, dass k Kugeln aus den M Kugeln gezogen werden und die restlichen $n - k$ Kugeln aus den $N - M$ Kugeln ohne die Eigenschaft.

$E(X) = n \cdot \frac{M}{N}$ $\qquad V(X) = n\frac{M}{N}\left(1 - \frac{M}{N}\right)\frac{N-n}{N-1}$

Beispiel:

Ein Studierender hat von 100 Aufgaben 70 bearbeitet. In einer Prüfung werden 5 dieser Aufgaben abgefragt. Mit welcher Wahrscheinlichkeit hat der Studierende genau 3 der Prüfungsaufgaben bereits bearbeitet?
Für die Zufallsvariable X, die die Anzahl der bereits bearbeiteten Aufgaben in der Prüfung angibt, gilt $X \sim Hyp(100, 70, 5)$.

Es gilt $P(X = 3) = \texttt{hygepdf}(3, 100, 70, 5) \approx 31.63\%$.

Poisson-Verteilung

Sie wissen, dass an einem Tag im einem bestimmten Skigebiet im Schnitt $\lambda = 4$ Unfälle passieren.

Es sei X die Zufallsvariable, die die Anzahl der Unfälle angibt.

Was ist $P(X = k)$?

Wir entwickeln nun dafür ein Modell.

- Man könnte das als $n = 24$-faches Bernoulli-Experiment modellieren, mit $p_n = \dfrac{1}{6} = \dfrac{\lambda}{n}$ als Wahrscheinlichkeit für Unfall (in einer Stunde $= \frac{1}{n}$ Tag).

  Dann ist der Erwartungswert gerade 4.

  Und wir hätten $P(X = k) = \binom{24}{k}\left(\frac{1}{6}\right)^k\left(1 - \frac{1}{6}\right)^{n-k}$

Probleme:

- Die Wahrscheinlichkeit für mehr als $n = 24$ Unfälle ist 0, was nicht realistisch ist.
- Bei Bernoulli-Experimenten tritt aber entweder Treffer oder nicht Treffer auf, und nie mehrere Treffer.
  Es könnten aber ja zwei Unfälle pro Stunde passieren.

Deshalb: Verfeinerung des Zeitintervalls, also Erhöhung von n.

- Man könnte das als $n = 48$-faches Bernoulli-Experiment modellieren, mit $p_n = \dfrac{4}{48} = \dfrac{\lambda}{n}$ als Wahrscheinlichkeit für Unfall (in einer halben Stunde $= \frac{1}{n}$ Tag).
  Dann ist der Erwartungswert auch 4.
  Und wir hätten $P(X = k) = \binom{n}{k}(p_n)^k(1 - p_n)^{n-k}$ mit $n = 48$, $p_n = \frac{1}{12}$.
  Hier haben wir dieselben Probleme wie oben (etwas abgeschwächt).

Idee: Wir lassen das Zeitintervall immer kleiner werden, also n immer grösser.

Poisson-Verteilung

Wir berechnen also $\lim\limits_{n\to\infty} \binom{n}{k} (p_n)^k (1-p_n)^{n-k}$.

$$\lim_{n\to\infty} \binom{n}{k} (p_n)^k (1-p_n)^{n-k} = \lim_{n\to\infty} \binom{n}{k} \left(\frac{\lambda}{n}\right)^k \left(1-\frac{\lambda}{n}\right)^{n-k} =$$

$$= \lim_{n\to\infty} \frac{n!}{k!(n-k)!} \left(\frac{\lambda}{n}\right)^k \left(1-\frac{\lambda}{n}\right)^{n-k} = \frac{\lambda^k}{k!} \lim_{n\to\infty} \frac{n!}{(n-k)!n^k} \left(1-\frac{\lambda}{n}\right)^n \left(1-\frac{\lambda}{n}\right)^{-k} =$$

$$= \frac{\lambda^k}{k!} \lim_{n\to\infty} \frac{n\cdot(n-1)\cdot\ldots\cdot(n-k+1)}{n^k} \left(1-\frac{\lambda}{n}\right)^n \left(1-\frac{\lambda}{n}\right)^{-k} =$$

$$= \frac{\lambda^k}{k!} \lim_{n\to\infty} \underbrace{\boxed{\frac{n}{n}}}_{\to 1} \cdot \underbrace{\boxed{\frac{n-1}{n}}}_{\to 1} \cdot \ldots \cdot \underbrace{\boxed{\frac{n-k+1}{n}}}_{\to 1} \cdot \underbrace{\boxed{\left(1-\frac{\lambda}{n}\right)^n}}_{\to e^{-\lambda}} \underbrace{\boxed{\left(1-\underbrace{\boxed{\frac{\lambda}{n}}}_{\to 0}\right)^{-k}}}_{\to 1} = \frac{\lambda^k}{k!} e^{-\lambda}$$

Beachte: k ist eine feste Zahl, die nicht von n abhängt.

Poisson-Verteilung

Definition:

Die Poisson-Verteilung kommt bei Zufallsvariablen zum Einsatz, welche die Anzahl Ereignisse einer bestimmten Art in einem Zeit- und/oder Ortsintervall beschreiben.

Falls im Mittel λ Ereignisse auftreten, dann ist X Poisson verteilt mit Parameter λ.

Schreibweise: $X \sim Poi(\lambda)$

Dichte von X: $f(k) = \frac{\lambda^k}{k!} e^{-\lambda}$, $k = 0, 1, 2, 3, \ldots$

$E(X) = \lambda$

$V(X) = \lambda$

> Matlab-Funktionen:
> Dichte: `poisspdf`(k, λ)
> Verteilungsfunktion: `poisscdf`(k, λ)

Beispiele für die Anwendung der Poisson-Verteilung:

- X Anzahl Druckfehler auf einer Seite eines Buches
- X Anzahl Unfälle an einem Wochenende in einem Skigebiet
- X Anzahl falsch gewählter Telefon-Nummern an einem Tag
- X Anzahl Erdbeben in einem Jahr in einer bestimmten Region

Der Parameter λ wird bei konkreten Anwendungen mittels Stichproben bestimmt.

Beispiel: X Anzahl Druckfehler auf einer Seite eines Buches

Die mittlere Anzahl Druckfehler pro Seite ist 0.4.

Dann ist $X \sim Poi(0.4)$ ein gutes Modell.

Wir erhalten etwa:

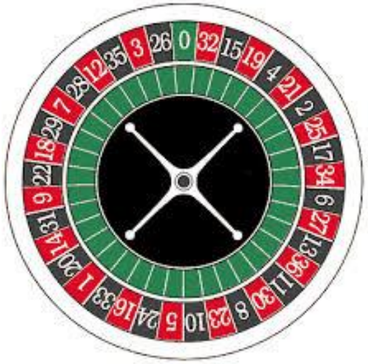
$P(X = 0) = e^{-\lambda} = poisspdf(0, 0.4) = 67.03\%$: kein Druckfehler auf der Seite

$P(X = 2) = \frac{\lambda^2}{2!} e^{-\lambda} = poisspdf(2, 0.4) = 5.36\%$: genau zwei Druckfehler auf der Seite

diskrete Verteilungen: Zusammenfassung

Alle Ereignisse gleichwahrscheinlich	Laplace
Treffer (WK p), nicht Treffer	$B(p)$
Anzahl Treffer(WK p) in n unabhängigen Versuchen	$Bin(n, p)$
Versuche bis erster Treffer (WK p) in unabhängigen Versuchen	$Geo(p)$
Anzahl günstiger Kugeln beim Ziehen ohne Zurücklegen von n Kugeln aus N Kugeln mit M günstigen Kugeln	$Hyp(N, M, n)$
Verteilung für "seltene Ereignisse" mit im Schnitt λ Ereignissen pro Zeit/Ort-Einheit	$Poi(\lambda)$

Beispiel: Wir spielen mit folgender Strategie Roulette:



Wir setzen Fr. 100 auf rot.

Wenn wir gewinnen, hören wir auf.

Andernfalls verdoppeln wir den Einsatz und beginnen von vorne.

In jedem Fall hören wir also mit einem Gewinn von Fr. 100 auf.
Was ist das Problem an der Strategie? Unsere beschränkten finanziellen Ressourcen:
Nehmen wir einmal an, wir hätten Fr. 50000 Franken zur Verfügung.
Was ist dann der erwartete Gewinn?
Wenn wir die ersten 8 Spiele verlieren, dann haben wir $100 + 2 \cdot 100 + 2^2 \cdot 100 + \ldots + 2^7 \cdot 100 = (2^8 - 1) \cdot 100 = 25500$ verloren.
Im nächsten Spiel müssten wir $2^8 \cdot 100 = 25600$ setzen, was zusammen mit unserem bisherigen Verlust mehr als 50000 wäre. Wir müssen also nach 8 verlorenen Spielen aufhören.
Die Zufallsvariable X, die den Gewinn/Verlust beschreibt, kann also die Werte 100 und -25500 annehmen.
Für den Erwartungswert erhalten wir:
$E(X) = -(19/37)^8 \cdot 25500 + 100 \cdot (1 - (19/37)^8) = -23.781 < 0.$
Auf lange Sicht wäre dies also keine einträgliche Strategie für den Spieler!
Auch wenn die Wahrscheinlichkeit Fr. 100 zu gewinnen bei mehr als 99.7% liegt.
(Aber wer gibt sich schon mit Fr. 100 zufrieden, wenn man mit Fr. 50000 ins Casino geht?)

Eigenschaften von Erwartungswert und Varianz

Satz: Es seien X, Y Zufallsvariablen und $a, c \in \mathbb{R}$.

Dann gilt (unter der Voraussetzung von Existenz im unendlichen Fall):

$$\left.\begin{array}{l} E(X+Y) = E(X) + E(Y) \\ E(aX) = aE(X) \end{array}\right\} \text{Der Erwartungswert ist linear}$$

$$E(X+c) = E(X) + c$$

$$V(X+c) = V(X)$$

$$V(aX) = a^2 V(X)$$

$$E(g(X)) = \sum_x g(x) P(X=x) \text{ für alle Funktionen } g : \mathbb{R} \to \mathbb{R}$$

Im Allgemeinen gilt $E(g(X)) \neq g(E(X))$.

Beispiel:

Wir werfen solange eine Münze, bis zum ersten Mal Kopf kommt, dann würfeln wir solange, bis eine 6 kommt.

Wie hoch ist die erwartete Anzahl an Schritten?

$X \sim Geo\left(\frac{1}{2}\right)$; $Y \sim Geo\left(\frac{1}{6}\right)$; $E(X+Y) = E(X) + E(Y) = 8$

Eigenschaften von Erwartungswert und Varianz

Beispiel:

Ein Computerhändler hat 3 Computer für \$500 das Stück gekauft, die er für \$1000 noch vor Weihnachten verkaufen will.
Der Hersteller hat eingewilligt, alle nicht verkauften Computer nach Weihnachten für \$200 zurückzunehmen

Es sei X die Zufallsvariable, die die Anzahl an verkauften Computern beschreibt.

Aus Erfahrung kennt der Händler die Verteilung von X:

x_i	0	1	2	3
$P(X = x_i)$	0.1	0.2	0.3	0.4

Wie hoch ist der erwartete Gewinn?

Der Gewinn ist: $h(X) = 1000 \cdot X - 1500 + (3 - X) \cdot 200 = 800X - 900.$

Also: $E(h(X)) = \underbrace{\boxed{-900}}_{= h(0)} \cdot 0.1 \underbrace{\boxed{-\ 100}}_{= h(1)} \cdot 0.2 + \underbrace{\boxed{700}}_{= h(2)} \cdot 0.3 + \underbrace{\boxed{1500}}_{= h(3)} \cdot 0.4 = 700.$

Serie 4, komplett