Spezielle Zufallsvariablen

Anleitung: Sie haben bis zum 7. Januar 2022, 23:55 Uhr, Zeit, Ihre Lösung für diese Aufgabe auf OLE einzureichen. Das heißt, Sie haben gut 2 Wochen regulärer Kurszeit. Denken Sie daran, dass Ihre Antworten auf jede Frage in die Endnote einfließen. Daher lohnt es sich, die Aufgabe einzureichen, auch wenn Sie sich nicht in der Lage fühlen, alle Fragen zu beantworten.

- Sie können Ihre Lösungen mit einem Textverarbeitungssystem (Word, Latex) oder per Hand ausarbeiten. Es ist möglicherweise einfacher, wenn Sie Ihre Antworten mit der Hand schreiben, da sie wahrscheinlich symbolische Rechnungen mit Brüchen, Potenzen und Integralen enthalten. Wenn Sie handschriftliche Lösungen einreichen, bemühen Sie sich, deutlich zu schreiben und Ihre Antworten so zu strukturieren und zu kommentieren, dass sie lesbar sind. Wenn Sie mit der Hand schreiben, reichen Sie die Antwort als gescanntes PDF-Dokument ein. (Reichen Sie kein Foto, sondern einen Scan ein, denn Fotos sind in der Regel nur schwer zu lesen).

- Erläutern Sie Ihre Antworten und den Ansatz, mit dem Sie sie erhalten haben. Erläutern Sie insbesondere, welche Regel Sie angewendet haben, wenn Sie ein bestimmtes Resultat der Vorlesung benutzt haben (z.B. „die Varianz einer Summe von unabhängigen Variablen ist die Summe der Varianzen"). Eine Kette von Formeln reicht nicht aus als Antwort.

- Jede Frage hat eine Gewichtung, die in Punkten ausgedrückt wird. Für Ihre Antwort erhalten Sie eine Note auf einer Skala von 0 bis 30. Die Gewichtung bestimmt, wie viel die Note zur Gesamtnote für die Hausarbeiten und damit zur Endnote beiträgt. Beachten Sie, dass wir bei der Berechnung der Endnote für jede Frage das Maximum aus der Note für die Frage und der Prüfungsnote bilden. Es ist daher kein großer Verlust, wenn Sie eine niedrige Note für eine Frage erhalten: (i) die Note für die Frage kann durch die Prüfungsnote ausgeglichen werden, und (ii) sie hat keinen Einfluss auf die Noten für die anderen Fragen.

- Ihre Arbeit soll Ihre eigene Leistung darstellen. Wir erwarten jedoch nicht von Ihnen, dass Sie allein arbeiten. Es ist in Ordnung, die Aufgaben zu besprechen und gemeinsam nach Lösungen zu suchen, aber jeder Student muss seine Lösungen separat aufschreiben und einreichen. Es gehört zum guten akademischen Standard, Mitarbeiter zu erwähnen. Wenn Sie also mit anderen zusammengearbeitet haben, geben Sie bitte deren Namen an.

- Wenn Sie diesen Kurs bereits im letzten Jahr belegt haben, müssen Sie nicht die gleichen Antworten wie im letzten Jahr einreichen, um Punkte zu erhalten. Reichen Sie nur dann Antworten auf eine Frage ein, wenn sie sich von den Antworten des letzten Jahres unterscheiden. Bei anderen Fragen geben Sie einfach an, dass Sie die gleiche Note wie im Vorjahr erhalten möchten. Das erspart uns die wiederholte Benotung der gleichen Arbeit.

1 Schlüpfende Küken

Angenommen, wir betreiben eine Hühnerfarm. Die Erfahrung zeigt, dass nicht aus jedem Ei ein Küken schlüpft, sondern dass dies mit einer Wahrscheinlichkeit p geschieht, wobei $0 < p < 1$ ist.
Auch überlebt nicht jedes geschlüpfte Küken bis zum Erwachsenenalter, sondern auch dies geschieht mit einer Wahrscheinlichkeit r, mit $0 < r < 1$.
Da das Ausbrüten der Eier und die Aufzucht der Küken unter gründlich überwachten Bedingungen geschehen, können wir annehmen, dass das Schlüpfen aus einem Ei unabhängig vom Schlüpfen aus den anderen ist, genauso, wie das Überleben eines Kükens unabhängig vom Überleben der anderen ist.
Derzeit befinden sich n Eier im Brutschrank.

1. Wie groß ist die Wahrscheinlichkeit, dass aus diesen n Eiern k Hühner schlüpfen?
2. Wie groß ist die Wahrscheinlichkeit, dass aus den n Eiern k Hühner hervorgehen, die bis zum Erwachsenenalter überleben?

Geben Sie für jede Frage eine Wahrscheinlichkeitsfunktion (= probability mass function) an, die den Fall modelliert, und nennen Sie den Namen der zugehörigen Verteilung und ihre Parameter.

(Gewicht: 8% dieser Arbeit)

2 Brechende Spaghetti

Nehmen wir an, wir lassen eine Spaghetti-Nudel fallen, so dass sie in zwei Teile zerbricht, und wiederholen dies viele Male. Eine Nudel hat die Länge l.
Wir nehmen an, dass die Positionen der Bruchstelle gleichmäßig verteilt sind. Wir interessieren uns für die Länge L des längeren Stücks.

1. Welches ist die Verteilungsfunktion (cumulative distribution function, cdf) von L? Welches die Dichte?
2. Was ist die durchschnittliche Länge des längeren Stücks?
3. Was ist die Varianz der Länge des längeren Stücks?

Hinweis: Nehmen Sie zunächst an, dass $l = 1$ ist.

(Gewicht: 10% dieser Arbeit)

3 Alice wartet

In einem Postamt gibt es 2 Angestellte. Alice betritt das Postamt, während 2 andere Kunden, Bob und Claire, von den beiden Angestellten bedient werden. Alice ist die nächste in der Schlange. Nehmen wir an, dass die Zeit, die ein Angestellter mit der Bedienung eines Kunden verbringt, *Exp*(λ)-verteilt ist.

1. Wie groß ist die Wahrscheinlichkeit, dass Alice als letzte der 3 Kunden bedient wird?
2. Wie groß ist der Erwartungswert für die gesamte Zeit, die Alice im Postamt verbringt?

Hinweis: Was wissen Sie über das Minimum von zwei exponentiell verteilten Variablen?

(Gewicht: 8% dieser Arbeit)

4 Wie schief ist die Exponentialverteilung? (How Skewed is the Exponential Distribution?)

Mittelwert und Varianz sind zwei Parameter, die die Form der Verteilung einer Zufallsvariablen beschreiben. Der Mittelwert ist der Schwerpunkt, während die Quadratwurzel der Varianz, die Standardabweichung, angibt, wie weit die Werte vom Schwerpunkt entfernt sind.
Der Mittelwert der Verteilung von $\mathcal{X}$ ist der Erwartungswert $\mu = E[\mathcal{X}]$ of $\mathcal{X}$. Um die Varianz zu ermitteln, normalisieren wir die Variable so, dass ihre Verteilung den Mittelwert 0 hat, was dadurch erreicht wird, dass wir μ von $\mathcal{X}$ subtrahieren. Die Varianz ist dann der Erwartungswert $E[(\mathcal{X}-\mu)^2]$ des *Quadrats* der normalisierten Version von $\mathcal{X}$.
Vom Statistiker Karl Pearson stammt die folgende Idee zur Messung der *Schiefe* (engl. *skewness*) oder *Wölbung* (engl. *kurtosis*) der Verteilung von $\mathcal{X}$. Wir normalisieren $\mathcal{X}$ zu einer Variablen, deren Verteilung die gleichen Proportionen hat wie die von $\mathcal{X}$, aber deren Mittelwert 0 und deren Varianz 1 ist. Dazu subtrahieren wir zunächst μ und teilen dann die Differenz durch die Standardabweichung σ, d.h. die Quadratwurzel der Varianz. Als Schiefe der Verteilung von $\mathcal{X}$ definieren wir dann den Erwartungswert der dritten Potenz der normierten Version von $\mathcal{X}$, d.h.

$$Kurt(\mathcal{X}) = E\left[\left(\frac{\mathcal{X}-\mu}{\sigma}\right)^3\right]. \tag{1}$$

Dahinter steckt die folgende Intuition:

- Ist die Verteilung symmetrisch um den Mittelwert, dann werden die dritten Potenzen der Werte links von μ, welche negativ sind, durch die dritten Potenzen der Werte rechts von μ ausgeglichen, so dass eine solche Verteilung die Schiefe 0 hat.
- Wenn man eine Zahl z mit $|z| < 1$ mit 3 potenziert, dann ist $|z^3| < |z|$, und wenn man ein z mit $|z| > 1$ mit 3 potenziert, dann ist $|z^3| > |z|$. Wenn man also $\frac{\mathcal{X}-\mu}{\sigma}$ mit 3 potenziert und den Erwartungswert nimmt, dann wird der Beitrag derjenigen

Werte von $\mathcal{X}$ verringert, deren Abstand zu μ kleiner als σ ist, und der Beitrag der Werte mit größerem Abstand wird verstärkt.

Auf diese Weise erhalten diejenigen Werte von $\mathcal{X}$ das meiste Gewicht, die weit von μ entfernt sind und nicht durch Werte auf der anderen Seite von μ ausgeglichen werden. Wenn es mehr solche Werte links von μ gibt, ist Ausdruck (1) negativ, wenn sie sich ausgleichen, ist er nahe 0, und wenn es mehr rechts davon gibt, ist er positiv.

Hier Ihre Aufgaben. Zur Erinnerung: Die Varianz von X erfüllt die Gleichung

$$\mathit{Var}(\mathcal{X}) = E\left[\mathcal{X}^2\right] - \mu^2.$$

1. Finden Sie eine Gleichung, in der $\mathit{Kurt}(\mathcal{X})$ ausgedrückt wird durch $E\left[\mathcal{X}^3\right]$ und eventuell μ und σ.

Wir nehmen an, $\mathcal{X}$ ist exponentialverteilt mit der Rate λ, d.h. $\mathcal{X} \sim Exp(\lambda)$.

2. Bestimmen Sie $E[\mathcal{X}^3]$.

   Hinweis: Verwenden Sie das LOTUS-Theorem und die Tatsache, dass $E[\mathcal{X}] = 1/\lambda$ und $\mathit{Var}(\mathcal{X}) = 1/\lambda^2$ ist. Sie können die Antwort mit einer einzigen Anwendung partieller Integration erhalten.

3. Bestimmen Sie $\mathit{Kurt}(\mathcal{X})$ mit Hilfe Ihrer Antwort auf die erste Teilfrage. Wie hängt $\mathit{Kurt}(\mathcal{X})$ von der Rate λ ab?

(Gewicht: 10% dieser Arbeit)

5 Bobs Emails

An Wochentagen (Montag-Freitag) und an Wochenenden (Samstag-Sonntag) erhält Bob eine unterschiedliche Anzahl von E-Mails. In beiden Fällen kann die Anzahl der eingehenden E-Mails durch eine Poisson-Verteilung modelliert werden, mit einer Rate von einer E-Mail alle 6 Minuten für die Wochentage und einer E-Mail alle 30 Minuten für die Wochenenden.

1. Wie groß ist die Wahrscheinlichkeit, dass Bob während einer bestimmten Stunde an einem Wochentag k E-Mails erhält?

2. Wie groß ist die Wahrscheinlichkeit, dass Bob während der vier Stunden von 13.00 bis 17.00 Uhr am Sonntag keine E-Mails erhält?

3. Wie ist die Verteilung der Anzahl der E-Mails, die Bob während eines Arbeitstages von 24 Stunden erhält? (Das heißt, wie groß ist die Wahrscheinlichkeit dafür, dass er 0, 1, 2, usw. E-Mails pro Tag erhält?)

4. Wie hoch ist die durchschnittliche Anzahl der E-Mails, die er an einem Tag erhält? Erläutern Sie Ihre Antwort.

5. Angenommen, wir wählen einen zufälligen Tag (unter der Woche oder am Wochenende) und ein zufälliges Intervall von einer Stunde Länge an diesem Tag. Bob erhält in diesem Intervall 5 Mails. Wie groß ist die Wahrscheinlichkeit dafür, dass der gewählte Tag ein Wochentag ist?

(Gewicht: 10% dieser Arbeit)

6 Wann sehen wir die erste Sechs?

Angenommen, Sie würfeln so lange, bis Sie zum ersten Mal eine Sechs erhalten.

1. Wie groß ist die Wahrscheinlichkeit, dass Sie genau k Würfe benötigen, mit $1 \leq k < \infty$, bis Sie zum ersten Mal eine Sechs sehen?

2. Verallgemeinern Sie Ihr Ergebnis auf den Fall, dass Sie k unabhängige Wiederholungen eines Bernoulli-Experiments mit der Verteilung $Bern(p)$ durchführen, bis sich der erste Erfolg einstellt.

3. Sei $\mathcal{X}$ die Anzahl der Wiederholungen bis zum ersten Erfolg. Wie groß ist der Erwartungswert $E[\mathcal{X}]$? Wie groß ist der Erwartungswert im Fall des Würfelexperiments zu Beginn?

(Gewicht: 20% dieser Arbeit)

7 Wie groß sind US-Frauen?

Die Körpergröße von erwachsenen Frauen in den Vereinigten Staaten ist mit einem Mittelwert von 164 cm und einer Standardabweichung von 6 cm verteilt. Wir nehmen an, dass die Größe normalverteilt ist. Ermitteln Sie die Wahrscheinlichkeit, dass eine eine zufällig ausgewählte Frau

1. weniger als 160cm groß ist;

2. weniger als 178 cm groß ist;

3. zwischen 160 und 178 cm groß ist.

4. Alice ist 182 cm groß. Wie viel Prozent der Frauen sind kleiner als Alice?

5. Ermitteln Sie die Wahrscheinlichkeit dafür, dass der Durchschnitt der Körpergröße von zwei zufällig ausgewählten Frauen 167 cm übersteigt.

6. Ermitteln Sie die entsprechende Wahrscheinlichkeit für vier zufällig ausgewählte Frauen.

(Gewicht: 12% dieser Arbeit)

8 Wer gewinnt beim Roulette?

Ein Rouletterad hat 37 Fächer, die mit den Zahlen 0 und 1 bis 36 nummeriert sind. Wenn Sie 1 auf eine bestimmte Zahl setzen, gewinnen Sie entweder 35, wenn die Roulettekugel auf dieser Zahl landet, oder Sie verlieren 1, wenn sie nicht auf dieser Zahl landet. Wenn Sie ständig solche Einsätze tätigen, schätzen Sie die Wahrscheinlichkeit, dass Sie

1. nach 35 Einsätzen einen Gewinn haben;
2. nach 1.000 Einsätzen einen Gewinn haben;
3. nach 100.000 Einsätzen einen Gewinn haben.

Beachten Sie, dass es hier nicht darum geht, zu schätzen, wie viel Sie gewinnen oder verlieren, sondern ob Sie am Ende mehr Geld haben als zu Beginn. Gehen Sie davon aus, dass bei jedem Wurf der Roulettekugel die Wahrscheinlichkeit gleich groß ist, auf einer der 37 Zahlen zu landen.

(Gewicht: 12% dieser Arbeit)

9 Wie falsch sind gerundete Zahlen?

Fünfzig Zahlen werden auf die nächste ganze Zahl gerundet und dann summiert. Wir nehmen an, die einzelnen Rundungsfehler sind gleichmäßig zwischen $-0,5$ und $0,5$ verteilt.

1. Wie groß ist angenähert die Wahrscheinlichkeit, dass die resultierende Summe um mehr als 3 von der exakten Summe abweicht?
2. Verallgemeinern Sie das vorige Ergebnis auf den Fall, dass es n Zahlen gibt und wir uns für Wahrscheinlichkeit interessieren, dass, für ein gegebenes $\epsilon > 0$, die sich ergebende Summe von der exakten Summe um mehr als ϵ abweicht.

Um dies zu präzisieren: wir fragen in Teil 2 nach der Wahrscheinlichkeit, dass der Absolutbetrag der Fehlersumme größer als ϵ ist. Teil 1 betrifft den Spezialfall $\epsilon = 3$.

Hinweis: Es kann sinnvoll sein, zunächst den allgemeinen Fall zu behandeln und dann den speziellen.

(Gewicht: 10% dieser Arbeit)