Musterlösung zu Übungsblatt 3

erstellt von Robin Dass
LaTeX-Vorlage: Jan Stricker

9. *Von der Multinomialverteilung zurück zur Binomialverteilung.*
a) Die Multinomialverteilung ist eine diskrete Wahrscheinlichkeitsverteilung und die Verallgemeinerung der Binomialverteilung. $X = (X_1, X_2, X_3)$ ist laut Angabe eine multinomialverteilte ZVe. Im ersten Teil dieser Aufgabe soll ohne rechnerischen Beweis die Verteilung von $X_1 + X_2$ dargelegt werden. Hierbei kann es hilfreich sein, sich einen "Würfel mit drei Seiten" vorzustellen, wobei X_j, $j = 1, 2, 3$ beim 10-maligen Würfeln die jeweiligen Ausgänge j zählt und somit $X_1 + X_2 + X_3 = 10$ ist. Zunächst definieren wir die ZVe

$$T := (T_1, ..., T_{10}),$$

um die Ausgänge festzuhalten. Daraus folgt unmittelbar, dass

$$X_j = \#\{i : 1 \le i \le 10, T_i = j\}.$$

Außerdem definieren wir

$$\hat{T}_i := \begin{cases} 1, & T_i \in \{1,2\} \\ 0, & \text{sonst} \end{cases}.$$

Das Zusammenfassen der beiden Ausgänge 1 und 2 zu einem Ausgang macht aus dem 10-maligen $(\frac{1}{3}, \frac{1}{4}, \frac{5}{12})$-Würfeln $(T_1, \dots, T_{10})$ einen 10-fachen $(\frac{1}{3} + \frac{1}{4})$-Münzwurf $(\hat{T}_1, \dots, \hat{T}_{10})$. Nun gilt

$$\#\{i : \hat{T}_i = 1\} = \#\{i : T_i = 1\} + \#\{i : T_i = 2\} = X_1 + X_2 =: Y.$$

Y zählt also wie oft $\hat{T}_i = 1$ ist. Es folgt

$$Y = \sum_{i=1}^{10} \hat{T}_i.$$

$Y = X_1 + X_2$ ist somit die Summe von Bernoulli-verteilten ZVe, also ist $X_1 + X_2$ $\mathrm{Bin}(10, \frac{1}{4} + \frac{1}{3})$-verteilt.

b) Im zweiten Teil dieser Aufgabe wird nun rechnerisch gezeigt werden, dass $X_1 + X_2$ binomialverteilt ist. Hierzu berechnen wir die Wahrscheinlichkeit des Ereignisses $\{X_1 + X_2 = k\}$ mittels Summation der Verteilungsgewichte aller Besetzungen $(b_1, b_2, b_3) \in S_{10,3}$ mit $b_1 + b_2 = k$. Die Summation über der (in VL 2b, F.9 behandelten) Gewichte der Multinomialverteilung liefert:

$$\begin{aligned}
&\mathbf{P}(\{X_1 + X_2 = k\}) \\
&\qquad = \sum_{b_1+b_2=k} \binom{n}{b_1, b_2, n-k} p_1^{b_1} p_2^{b_2} p_3^{n-k} \\
&\qquad = \sum_{b_1+b_2=k} \frac{n!}{b_1! b_2! (n-k)!} p_1^{b_1} p_2^{b_2} p_3^{n-k}.
\end{aligned}$$

Mit $b_1 = k - b_2$ rechnen wir $\binom{n}{k}\binom{k}{b_2} = \frac{n!}{k!(n-k)!}\frac{k!}{b_2!(k-b_2)!} = \frac{n!}{(n-k)!b_2!}\,\frac{1}{b_1!}$ und erhalten damit unter Verwendung des binomischen Lehrsatzes die Gewichte der Binomialverteilung:

$$\sum_{b_1+b_2=k} \frac{n!}{b_1!b_2!(n-k)!} p_1^{b_1} p_2^{b_2} p_3^{n-k} \quad = \textstyle\sum_{b_2=0}^{k} \binom{n}{k}\binom{k}{b_2} p_1^{k-b_2} p_2^{b_2} p_3^{n-k} \quad \overset{\text{binom. Lehrsatz}}{=} \binom{n}{k}(p_1+p_2)^k p_3^{n-k}.$$

10.S *Wahrscheinlichkeit für "genau pari-pari".*
a) In dieser Aufgabe liegt ein $3n$-faches $(\frac{1}{3}, \frac{1}{3}, \frac{1}{3})$-Würfeln vor und auch hier zählt (wie in Aufgabe 9) X_j, $j = 1, 2, 3$, die Anzahl der geworfenen j. Zur Berechnung von $\mathbf{P}((X_1, X_2, X_3) = (n, n, n))$ verwenden wir ebenfalls die Gewichte der Multinomialverteilung.

Allgemein gilt für eine multinomialverteilte ZVe $Y = (Y_1, ..., Y_g)$:

$$\mathbf{P}((Y_1, ..., Y_g) = (b_1, ..., b_g)) = \binom{n}{b_1, ..., b_g} p_1^{b_1} \cdot ... \cdot p_g^{b_g}.$$

i) für $n = 2$:

$$\mathbf{P}((X_1, X_2, X_3) = (2, 2, 2)) = \binom{3 \cdot 2}{2, 2, 2} \left(\frac{1}{3}\right)^2 \left(\frac{1}{3}\right)^2 \left(\frac{1}{3}\right)^2 = \frac{10}{81} \approx 12.35\%$$

und für $n = 3$:

$$\mathbf{P}((X_1, X_2, X_3) = (3, 3, 3)) = \binom{3 \cdot 3}{3, 3, 3} \left(\frac{1}{3}\right)^3 \left(\frac{1}{3}\right)^3 \left(\frac{1}{3}\right)^3 = \frac{560}{6561} \approx 8.54\%$$

ii) Unter Verwendung der Stirling-Formel $n! \sim \sqrt{2\pi n}\left(\frac{n}{e}\right)^n$ (VL 1b, F.11) und allgemein gewähltem $n \in \mathbb{N}$ rechnen wir:

$$\mathbf{P}((X_1, X_2, X_3) = (n, n, n)) = \binom{3 \cdot n}{n, n, n} \left(\frac{1}{3}\right)^n \left(\frac{1}{3}\right)^n \left(\frac{1}{3}\right)^n$$

$$= \frac{(3n)!}{n!\; n!\; n!} \left(\frac{1}{3}\right)^n \left(\frac{1}{3}\right)^n \left(\frac{1}{3}\right)^n$$

$$\overset{\text{Stirling}}{\approx} \frac{\sqrt{2\pi 3n}\left(\frac{3n}{e}\right)^{3n}}{\left(\sqrt{2\pi n}\left(\frac{n}{e}\right)^n\right)^3} \left(\frac{1}{3}\right)^{3n}$$

$$= \frac{\sqrt{6\pi n}\; 3^{3n} \left(\frac{n}{e}\right)^{3n}}{(\sqrt{2\pi n})^3 \left(\frac{n}{e}\right)^{3n}} \left(\frac{1}{3}\right)^{3n}$$

$$= \frac{\sqrt{6\pi n}}{2\pi n\; \sqrt{2\pi n}} \left(\frac{3}{3}\right)^{3n}$$

$$= \frac{\sqrt{3}\sqrt{2\pi n}}{2\pi n\sqrt{2\pi n}}$$

$$= \frac{\sqrt{3}}{2\pi n}.$$

Nun setzten wir $n = 2$ und $n = 3$ in die soeben hergeleitete Näherung für das Multinomialgewicht von (n, n, n) ein:
für $n = 2$:

$$\mathbf{P}((X_1, X_2, X_3) = (2, 2, 2)) \overset{\text{Stirling}}{\approx} \frac{\sqrt{3}}{2\pi \cdot 2} \approx 13.78\%$$

für $n = 3$:

$$\mathbf{P}((X_1, X_2, X_3) = (3, 3, 3)) \overset{\text{Stirling}}{\approx} \frac{\sqrt{3}}{2\pi \cdot 3} \approx 9.19\%.$$

Die Approximation mit der Stirling-Formel liefert in i) und ii) verglichen mit den exakten Werten einen relativen Fehler von ca. 10%. Auch hier sieht man: die Stirling-Formel ist nicht nur eine gute asymptotische Näherung für große n, sondern auch schon brauchbar für kleine n.

b) Wir starten das im Moodle-Kurs bereitgestellte R-Programm *deFinettiDreieck.R* und setzen $n = 9$ und $k = 1000$. Dabei ist n die Anzahl der Kugeln, welche in 3 gleich großen Kisten zufällig geworfen werden und k die Anzahl der Wiederholungen am Ende. Folglich wird angegeben wie die Kugeln die drei Kisten besetzen. Dementsprechend wird dann das De-Finetti-Dreieck für die vorangehende Besetzung dargestellt. Anschließend wird dieser Vorgang 1000-mal wiederholt und es zeigt sich, dass die Angabe der Häufigkeit der Besetzung im Zentrum des De-Finetti-Dreiecks am größten ist, also gerade bei $(3, 3, 3)$ (siehe Abbildung 1, Anhang). Diese Häufigkeit spiegelt auch in etwa das Ergebnis aus Teil a) wider, denn bei unseren 1000 Durchläufen war die Trefferquote der exakt ausgeglichene Besetzung der drei Kisten durch die neun Kugeln $\frac{92}{1000} = 9,2\%$. (Übrigens war das ähnlich nahe beim exakt berechneten Wert des Multinomialgewichts wie unsere Näherung mit der Stirling-Formel. Wenn Ihr Rechner stark genug und/oder Sie geduldig sind, können Sie das Monte-Carlo Experiment ja auch noch mit $k = 10000$ durchführen und sehen, wie nahe dann die relative Hufigkeit der Treffer von $(3, 3, 3)$ bei dem in Teil a) i) berechneten exakten Wert liegt.)

11. *Erwartungswert beim zufälligen Ziehen - mit und ohne Verzerrung.*
a) i) Hier müssen wir den Erwartungswert einer auf $\{1, ..., g\}$ uniform verteilten ZVen U berechnen:

$$\mathbf{E}[U] = \sum_{i=1}^{g} \mathbf{P}(U = i) \cdot i = \sum_{i=1}^{g} \frac{1}{g} \cdot i = \frac{1}{g} \sum_{i=1}^{g} i \overset{\text{kleiner Gauß}}{=} \frac{1}{g} \frac{g(g+1)}{2} = \frac{g+1}{2}.$$

ii) In diesem Aufgabenteil ist X eine $\{1, ..., g\}$-wertige ZVe. Um den Erwartungswert von X zu berechnen, gehen wir wie folgt vor:

$$\mathbf{E}[X] = \sum_{a \in \{1,...,g\}} a\mathbf{P}(X = a) = \sum_{a \in \{1,...,g\}} a\rho(a). \tag{1}$$

Nun wissen wir nach Aufgabenstellung, dass $\rho(a)$ proportional zu $\frac{1}{a}$ ist. Also gibt es eine Konstante $c \in \mathbb{R}$ mit $\rho(a) = c\frac{1}{a}$ ist. Außerdem gilt, dass die Summe aller Wahrscheinlichkeitsgewichte 1 ergibt:

$$1 = \sum_{a \in \{1,...,g\}} \rho(a) = \sum_{a \in \{1,...,g\}} c\frac{1}{a} = 1c + \frac{1}{2}c + \frac{1}{3}c + ... + \frac{1}{g}c = c(1 + \frac{1}{2} + \frac{1}{3} + ... + \frac{1}{g})$$

$$\Rightarrow c = \frac{1}{1 + \frac{1}{2} + \frac{1}{3} + ... + \frac{1}{g}} = \frac{1}{h_g} \text{ und } \rho(a) = \frac{1}{h_g a}.$$

Mit (1) folgt dann:

$$\mathbf{E}[X] = \sum_{a \in \{1,\ldots,g\}} a \frac{1}{h_g a} = \frac{g}{h_g}.$$

b) i) Sei $X =$ "Höhe des Baums", dann ist der Erwartungswert von X:

$$\mathbf{E}[X] = 40\% \cdot 1 + 30\% \cdot 2 + 30\% \cdot 3 = \frac{19}{10}.$$

ii) Zum Aufwärmen betrachten wir eine Situation, die anders ist als die in der Aufgabe beschriebene, und zwar mit $g = 3$, und je einem Baum der Höhe 1, 2 bzw. 3. Nach Voraussetzung ist die Wahrscheinlichkeit einen Baum zu wählen proportional zur Höhe des Baums. So muss bspw. "der Baum der Höhe 2", doppelt so wahrscheinlich hinsichtlich der Wahl sein, wie "der Baum der Höhe 1". Kurzes Überlegen und Probieren (beachte $1 + 2 + 3 = 6$) liefert: $\mathbf{P}(X = 1) = \frac{1}{6}$, $\mathbf{P}(X = 2) = \frac{2}{6}$, $\mathbf{P}(X = 3) = \frac{3}{6}$. Auch hier lässt sich der Erwartungswert nun analog zur b) i) berechnen:

$$\mathbf{E}[X] = \frac{3}{6} \cdot 3 + \frac{2}{6} \cdot 2 + \frac{1}{6} \cdot 1 = \frac{7}{3}.$$

Jetzt wenden wir uns der tatsächlichen Aufgabenstellung zu: Die Wahrscheinlichkeit, (irgend)einen Baum der Höhe b zu ziehen, ist proportional zu b **und** zum Anteil der Bäume mit Höhe b, also

$$\rho(1) = c \cdot 4 \cdot 1, \quad \rho(2) = c \cdot 3 \cdot 2, \quad \rho(3) = c \cdot 3 \cdot 3.$$

Wegen $4 + 6 + 9 = 19$ folgt

$$\rho(1) = \frac{4}{19}, \quad \rho(2) = \frac{6}{19}, \quad \rho(3) = \frac{9}{19}.$$

Eine Zufallsvariable H mit *diesen* Verteilungsgewichten hat den Erwartungswert

$$\mathbf{E}[H] = 1 \cdot \frac{4}{19} + 2 \cdot \frac{6}{19} + 3 \cdot \frac{9}{19} = \frac{43}{19} \approx 2\frac{1}{4}.$$

Das ist der gefragte Erwartungswert (gemessen in Metern).

12.S *Der Charme der Austauschbarkeit.*
a) i) Hier zeigen wir, dass für jedes $j \in \{1, ..., 4n\}$ gilt: $\mathbf{P}(X_2 = j) = \frac{1}{4n}$. Da $(X_1, \ldots, X_{4n})$ eine rein zufällige Permutation ist, arbeiten wir nach dem Prinzip "Anzahl günstige durch Anzahl mögliche". Wir zählen die Permutationen $(a_1, \ldots, a_{4n})$ von $1, \ldots, 4n$ ab, für die $a_2 = i$ gilt. Für a_1 gibt es $4n - 1$ Möglichkeiten, dann für a_3 noch $4n - 2$, usw., insgesamt also $(4n - 1)!$

Da es insgesamt $(4n)!$ Permutationen von $1, ..., 4n$ gibt, erhalten wir nach dem oben angegebenem Prinzip:

$$\mathbf{P}(X_2 = j) = \frac{(4n-1)!}{(4n)!} = \frac{1}{4n}, \quad j = 1, \ldots, 4n.$$

Damit ist X_2 auf der Menge $\{1, ..., 4n\}$ uniform verteilt.

ii) Mit genau demselben Argument wie in i) folgt:
$(*)$ Für jedes $i = 1, \ldots, 4n$ ist X_i uniform verteilt auf $\{1, \ldots, 4n\}$.
Um den Erwartungswert der Summe der Indikatorvariablen zu berechnen, nutzen wir einerseits die Linearität des Erwartungswertes und andererseits die Beziehung $(**)$ $\mathbf{E}[\mathbb{1}_E] = \mathbf{P}(E)$ aus:

$$\mathbf{E}\left[\sum_{i=1}^{3} \mathbb{1}_{\{X_i \le n\}}\right] \overset{\text{Linearität}}{=} \sum_{i=1}^{3} \mathbf{E}\left[\mathbb{1}_{\{X_i \le n\}}\right] \overset{(**)}{=} \sum_{i=1}^{3} \mathbf{P}(\{X_i \le n\}) \overset{(***)}{=} \sum_{i=1}^{3} \frac{1}{4} = \frac{3}{4},$$

wobei $(***)$ $\quad \mathbf{P}(X_i \le n) = \sum_{j=1}^{n} \mathbf{P}(X_i = j) \overset{(*)}{=} \sum_{j=1}^{n} \frac{1}{4n} = \frac{n}{4n} = \frac{1}{4}.$

b) i) Wir denken uns die 40 Kugeln nummeriert mit den Zahlen $1, \ldots, 40$, wobei die roten Kugeln die Nummern $1, \ldots, 10$ tragen, und ziehen in Gedanken alle 40 Kugeln. Die Folge der gezogenen Nummern ist dann eine rein zufällige Permutation von $1, \ldots, 40$, und wir sind genau im Fahrwasser von Aufgabenteil a) mit $n = 10$. Unter Berücksichtigung von a) i) gilt für das Ereignis $E :=$ “zweite gezogene Kugel ist rot”, dass die Wahrscheinlichkeit

$$\mathbf{P}(E) = \frac{10}{40} = \frac{1}{4}$$

ist.

ii) Analog zur b) i) definieren wir uns das Ereignis $F :=$ ”dritte gezogene Kugel ist blau” und bekommen “aus Symmetriegründen”

$$\mathbf{P}(F) = \frac{10}{40} = \frac{1}{4}.$$

iii) Der Erwartungswert einer Farbe ist nicht definiert, da Erwartungswerte nur für reellwertige ZVe definiert sind und eine Farbe (außer bei passender, dann aber noch zu definierender Codierung) keine reelle Zahl ist.

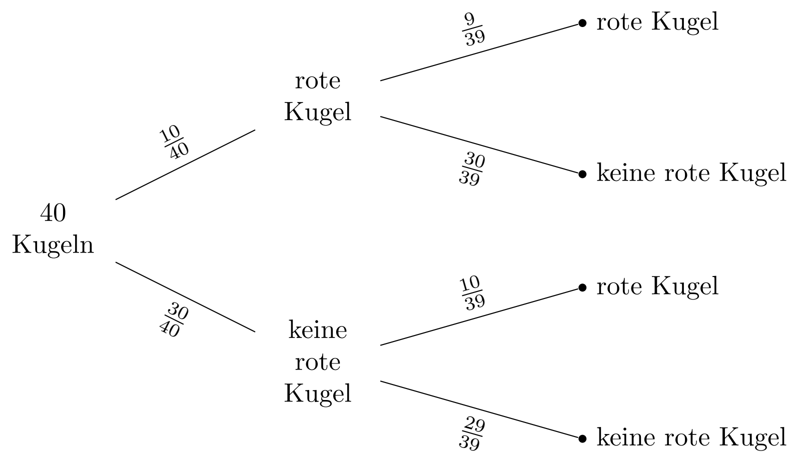
iv) ”Was ist der Erwartungswert der Anzahl der roten Kugeln in den ersten 3 Zügen?”
Um diese Frage zu beantworten, definieren wir das Ereignis $A_i :=$ ”im i-ten Zug eine rote Kugel ziehen”. Mithilfe der Indikatorvariablen erhalten wir:

$$\mathbf{E}\left[\sum_{i=1}^{3} \mathbb{1}_{A_i}\right] = \sum_{i=1}^{3} \mathbf{P}(A_i) = 3 \cdot \frac{1}{4} = \frac{3}{4}.$$

Genauso können wir das Ergebnis auch aus a) ii) ablesen, wieder mit $n = 10$.

Würden wir das Experiment (à la Monte Carlo) oft wiederholen, dann würden wir im Schnitt $\frac{3}{4}$ rote Kugeln in den ersten 3 Zügen zu sehen bekommen. Auf diese Interpretation des Erwartungswertes werden wir bei der Herleitung und Diskussion des *Gesetzes der Großen Zahlen* zurückkommen, vgl. dazu auch VL 3a Teil 5.

Zusatzbemerkung. Die folgende Argumentation zur Beantwortung von b) i) mithilfe eines Baumdiagrammes über die bedingten Wahrscheinlichkeiten (und die “Zerlegung nach dem ersten Schritt”) in vielen Köpfen fest verdrahtet, weil man sich am Anfang schwer vorstellen kann, dass die Frage beantworten kann, ohne das Ergebnis des ersten Zuges zu berücksichtigen.[1]



(Quelle: https://texample.net/tikz/examples/probability-tree/)

Gehen wir nun entlang der Pfade die Möglichkeiten durch beim zweiten Zug eine rote Kugel zu ziehen, wird das obige Ergebnis ersichtlich, denn

$$\mathbf{P}(E) = \frac{10}{40} \cdot \frac{9}{39} + \frac{30}{40} \cdot \frac{10}{39} = \frac{1}{4}.$$

[1] Wie wir gesehen haben, kann man diese Argumentation hier vermeiden, indem man den Charme der rein zufälligen Permutationen nutzt.

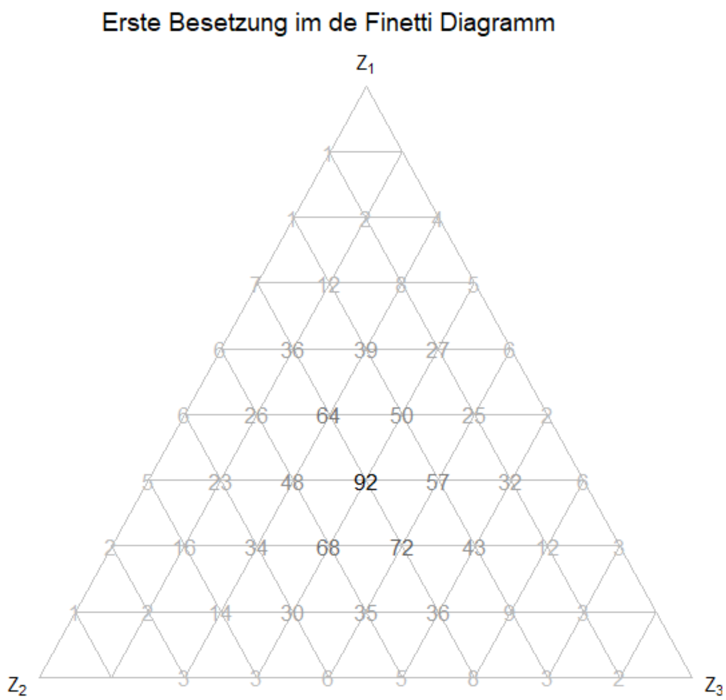


Abbildung 1: De-Finetti-Dreieck, $n = 9$, $k = 1000$