§13 Diskrete Verteilungen

▼

```
DiscreteEmpiricalPDF[daten_, z_] := Module[{n, min, max, uni},
    n = Length[daten];
    min = Min[daten];
    max = Max[daten];
    uni = Union[daten];
    If[MemberQ[daten, z], BinCounts[daten, {min, max + 1}][[Position[uni, z][[1, 1]]]] / n, 0] // N]
```

Wir haben bereits erwähnt, dass in der Verteilung $\mathbb{P}_Z$ einer Zufallsvariablen Z die gesamte Information über ihr zufälliges Verhalten enthalten ist. Damit ist die Verteilung $\mathbb{P}_Z$ einer Zufallsvariablen Z einer der fundamentalen Begriffe der Stochastik. Nun sind aber Verteilungen von Zufallsvariablen relativ komplizierte mathematische Gebilde, nämlich Abbildungen, die jeder (vernünftigen) Menge $B \subseteq \mathbb{R}$ eine reelle Zahl zuordnen und dabei die in Satz 12.2.2 angeführten Eigenschaften besitzen.

Will man mit Verteilungen - also mit W-Maßen $\mathbb{P}$ auf $\mathbb{R}$ - rechnen, so muss man mit diesen Verteilungen effizient arbeiten können. Dazu ist beispielsweise notwendig, eine Verteilung leicht und übersichtlich angeben und für beliebige Mengen $B \subseteq \mathbb{R}$ die Wahrscheinlichkeit $\mathbb{P}[B]$ leicht berechnen zu können. Wir werden uns in diesem Abschnitt mit der Frage befassen, wie dies bei diskreten Verteilungen geschehen kann.

13.1 Verteilungsdichten diskreter Verteilungen

Wir beginnen mit einer zentralen Eigenschaft diskreter Verteilungen $\mathbb{P}$ (vergleiche dazu Beispiel 12.2.3):

13.1.1 Bemerkung: Ist $\mathbb{P}$ eine diskrete Verteilung mit dem Träger $\mathbb{T}_\mathbb{P} \subseteq \mathbb{R}$, so gilt für alle $B \subseteq \mathbb{R}$

$$\mathbb{P}[B] = \mathbb{P}[B \cap \mathbb{T}_\mathbb{P}] = \sum_{z \in B \cap \mathbb{T}_\mathbb{P}} \mathbb{P}[\{z\}]$$

Die diskrete Verteilung $\mathbb{P}$ ist somit durch die Wahrscheinlichkeiten $\mathbb{P}[\{z\}]$ mit $z \in \mathbb{T}_\mathbb{P}$ vollständig bestimmt.

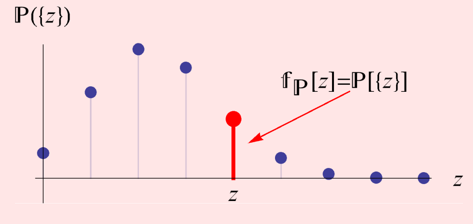
Wir nehmen diese Erkenntnis zum Anlass und definieren:

13.1.2 Definition: Unter der Verteilungsdichte (Probability Density Function oder kurz PDF) $\mathfrak{f}_\mathbb{P}$ einer diskreten Verteilung $\mathbb{P}$ mit dem Träger $\mathbb{T}_\mathbb{P} \subseteq \mathbb{R}$ versteht man die Abbildung

$$\mathfrak{f}_\mathbb{P} : \mathbb{R} \to \mathbb{R} \quad \text{mit} \quad \mathfrak{f}_\mathbb{P}[z] = \begin{cases} \mathbb{P}[\{z\}] & \text{für } z \in \mathbb{T}_\mathbb{P} \\ 0 & \text{sonst} \end{cases}$$

Wegen Bemerkung 13.1.1 ist die diskrete Verteilung $\mathbb{P}$ durch ihre Verteilungsdichte $\mathfrak{f}_\mathbb{P}$ vollständig bestimmt.

Die folgende Zeichnung zeigt die Verteilungsdichte $\mathfrak{f}_\mathbb{P}$ einer diskreten Verteilung $\mathbb{P}$. Die Punkte an den Stellen $z \in \mathbb{T}_\mathbb{P}$ entsprechen dabei den Wahrscheinlichkeiten $\mathfrak{f}_\mathbb{P}[z] = \mathbb{P}[\{z\}]$.



13.1.3 Bemerkung: Für die Verteilungsdichte $\mathbb{f}_{\mathbb{P}}$ einer diskreten Verteilung $\mathbb{P}$ gilt

a) Für alle $z \in \mathbb{R}$ ist $\mathbb{f}_{\mathbb{P}}[z] = \mathbb{P}[\{z\}] \geq 0$;

b) Die Summe über alle Werte $\mathbb{f}_{\mathbb{P}}[z]$ mit $z \in \mathbb{T}_{\mathbb{P}}$ ergibt den Wert 1, es gilt also

$$\sum_{z \in \mathbb{T}_{\mathbb{P}}} \mathbb{f}_{\mathbb{P}}[z] = 1$$

Diese beiden Eigenschaften sind charakteristisch in dem Sinn, dass jede Funktion $\mathbb{f}:\mathbb{R}\to\mathbb{R}$ mit diesen Eigenschaften als Verteilungsdichte einer diskreten Verteilung $\mathbb{P}$ aufgefasst werden kann. Man nennt diese beiden Eigenschaften daher die charakteristischen Eigenschaften der Verteilungsdichte einer diskreten Verteilung.

13.2 Verteilungsdichten in *Mathematica*

Mit dem folgenden Befehl lässt sich die Verteilungsdichte $\mathbb{f}$ einer in *Mathematica* implementierten diskreten Verteilungen (numerisch und formelmäßig) auswerten:

■ PDF[*distribution*, *z*]

wertet die Verteilungsdichte $\mathbb{f}$ der diskreten Verteilung *distribution* an der Stelle z aus. Man beachte, dass die Verteilungsdichte $\mathbb{f}$ einer diskreten Verteilung keine reelle Funktion im Sinn von *Mathematica* ist und sich damit auch nicht! mit Hilfe von Plot zeichnen lässt. Will man die Verteilungsdichte $\mathbb{f}$ einer diskreten Verteilung graphisch veranschaulichen, so verwende man dazu den Befehl ListPlot.

Wir wollen nun die Verteilungsdichten $\mathbb{f}$ der wichtigsten in *Mathematica* implementierten diskreten Verteilungen analysieren. Unser Ziel besteht darin, diese Verteilungsdichten formelmäßig explizit anzugeben und mit Hilfe von ListPlot und Manipulate auf dynamische Weise graphisch darzustellen.

13.2.1 Beispiel: Man ermittle die Verteilungsdichte der diskreten Gleichverteilung $\mathcal{U}[\{m, n\}]$ mit den Parametern $m < n \in \mathbb{Z}$ und zeichne diese Verteilungsdichte.

▼

Lösung: Wegen

```
DistributionDomain[DiscreteUniformDistribution[{m, n}]]
PDF[DiscreteUniformDistribution[{m, n}], z]
```

```
Range[m, n]
```

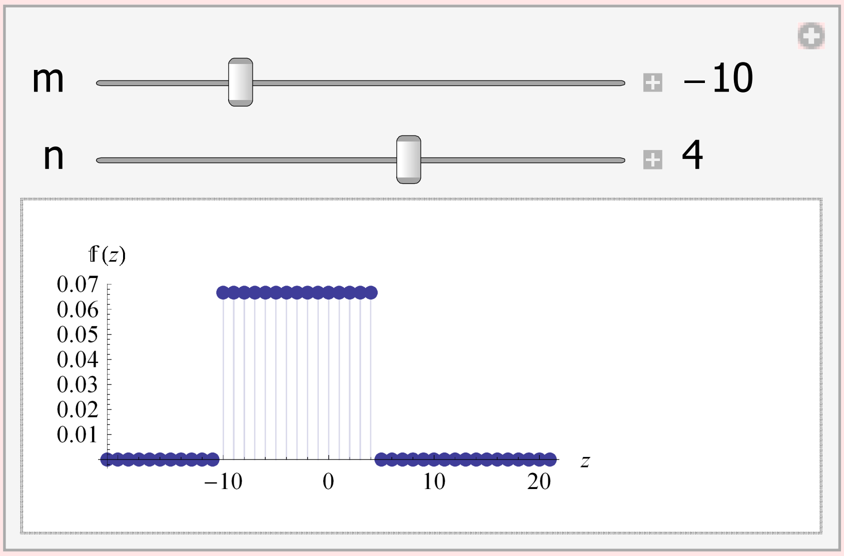
$$\frac{1}{1-m+n}$$

besitzt die diskrete Gleichverteilung $\mathcal{U}[\{m, n\}]$ die Verteilungsdichte

$$\mathbb{f}[z] = \begin{cases} 1/(n-m+1) & \text{für } z \in \{m, m+1, \ldots, n\} \\ 0 & \text{sonst} \end{cases}$$

Mit Hilfe von Manipulate lässt sich die Verteilungsdichte der diskreten Gleichverteilung $\mathcal{U}[\{m, n\}]$ für beliebige Werte von m und n auf dynamische Weise graphisch darstellen (wobei noch die Abfrage $m \leq n$ eingebaut wurde):

```
Manipulate[If[m ≤ n, ListPlot[Table[{z, PDF[DiscreteUniformDistribution[{m, n}], z]}, {z, −21, 21}],
      PlotStyle → PointSize[0.03], Filling → Axis, AspectRatio → 0.4, AxesOrigin → {−21, 0},
      PlotRange → All, AxesLabel → {z, 𝕗[z]}, ImageSize → {200, 100}], "m>n"],
  {m, −20, 20, 1, Appearance → "Labeled"}, {n, −20, 20, 1, Appearance → "Labeled"}]
```



13.2.2 Beispiel: Man ermittle die Verteilungsdichte der Binomialverteilung $\mathcal{B}[n, p]$ mit den Parametern $n \in \mathbb{N}$ und $0 \leq p \leq 1$ und zeichne diese Verteilungsdichte.

▼

Lösung: Wegen

```
DistributionDomain[BinomialDistribution[n, p]]
PDF[BinomialDistribution[n, p], z]
```

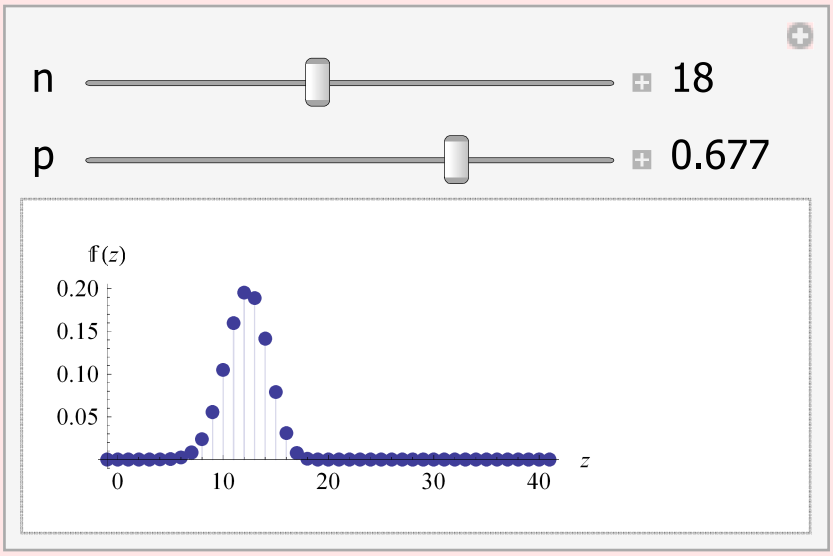
```
Range[0, n]

(1 - p)^(n-z) p^z Binomial[n, z]
```

besitzt die Binomialverteilung $\mathcal{B}[n, p]$ die Verteilungsdichte

$$\mathbb{f}[z] = \begin{cases} \binom{n}{z} p^z (1-p)^{n-z} & \text{für } z \in \{0, 1, 2, \ldots, n\} \\ 0 & \text{sonst} \end{cases}$$

Mit Hilfe von Manipulate lässt sich die Verteilungsdichte der Binomialverteilung $\mathcal{B}[n, p]$ für beliebige Werte von n und p auf dynamische Weise graphisch darstellen:

```
Manipulate[ListPlot[Table[{z, PDF[BinomialDistribution[n, p], z]}, {z, -1, 41}],
    PlotStyle → PointSize[0.03], Filling → Axis, AspectRatio → 0.4, AxesOrigin → {-1, 0}, PlotRange → All,
    AxesLabel → {z, f[z]}, ImageSize → {200, 100}],
  {n, 1, 40, 1, Appearance → "Labeled"}, {p, 0.1, 0.9, Appearance → "Labeled"}]
```



13.2.3 Beispiel: Man ermittle die Verteilungsdichte der negativen Binomialverteilung $\mathcal{NB}[n, p]$ mit den Parametern $n \in \mathbb{N}$ und $0 \leq p \leq 1$ und zeichne diese Verteilungsdichte.

▼

13.2.4 Beispiel: Man ermittle die Verteilungsdichte der hypergeometrischen Verteilung $\mathcal{H}[n, n_{suc}, n_{tot}]$ mit den Parametern $n, n_{suc}, n_{tot} \in \mathbb{N}$ und $\max[n, n_{suc}] \leq n_{tot}$ und zeichne diese Verteilungsdichte.

▼

13.2.5 Beispiel: Man ermittle die Verteilungsdichte der Poissonverteilung $\mathcal{P}[\lambda]$ mit dem Parameter $\lambda > 0$ und zeichne diese Verteilungsdichte.

▼

Lösung: Wegen

```
DistributionDomain[PoissonDistribution[λ]]
PDF[PoissonDistribution[λ], z]
```

```
Range[0, ∞]
```

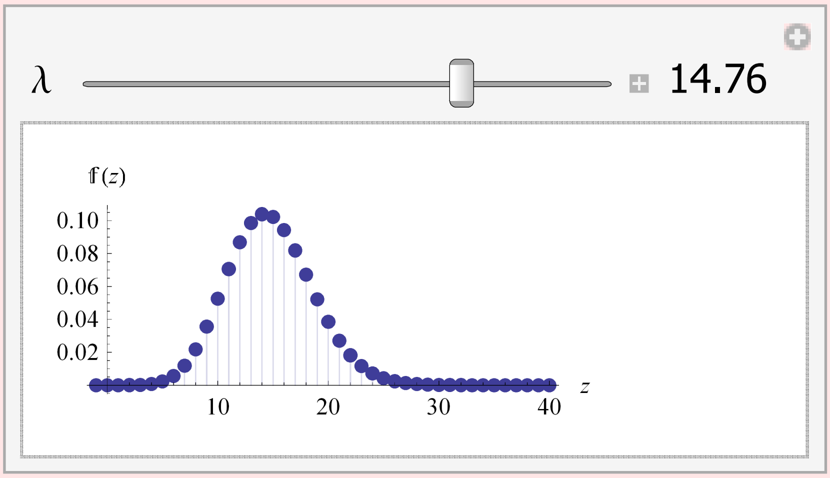
$$\frac{e^{-\lambda}\,\lambda^{z}}{z\,!}$$

besitzt die Poissonverteilung $\mathcal{P}[\lambda]$ die Verteilungsdichte

$$\mathbb{f}[z] = \begin{cases} e^{-\lambda}\,\lambda^{z}/z! & \text{für } z \in \{0, 1, 2, \ldots\} \\ 0 & \text{sonst} \end{cases}$$

Mit Hilfe von Manipulate lässt sich die Verteilungsdichte der Poissonverteilung $\mathcal{P}[\lambda]$ für beliebige Werte von λ auf dynamische Weise graphisch darstellen:

```
Manipulate[ListPlot[Table[{z, PDF[PoissonDistribution[λ], z]}, {z, −1, 40}],
    PlotStyle –> PointSize[0.03], Filling –> Axis, AspectRatio → 0.4, AxesOrigin → {0, 0}, PlotRange → All,
    AxesLabel → {z, f[z]}, ImageSize → {200, 100}],
  {λ, 0.01, 20, Appearance → "Labeled"}]
```



13.3 Verteilungsdichten diskreter Zufallsvariablen

Bisher haben wir uns mit den Verteilungsdichten $\mathbb{f}_{\mathbb{P}}$ von diskreten Verteilungen $\mathbb{P}$ beschäftigt. Nun wollen wir uns mit den Verteilungsdichten $\mathbb{f}_Z$ von diskreten Zufallsvariablen Z befassen.

13.3.1 Definition: Unter der Verteilungsdichte $\mathbb{f}_Z$ einer diskreten Zufallsvariablen Z versteht man die Verteilungsdichte der Verteilung $\mathbb{P}_Z$ von Z, also die Abbildung

$$\mathbb{f}_Z : \mathbb{R} \to \mathbb{R} \quad \text{mit} \quad \mathbb{f}_Z[z] = \begin{cases} \mathbb{P}[\{Z = z\}] & \text{für } z \in \mathbb{T}_Z \\ 0 & \text{sonst} \end{cases}$$

Wegen Bemerkung 13.1.1 ist die Verteilung $\mathbb{P}_Z$ der diskreten Zufallsvariablen Z durch ihre Verteilungsdichte $\mathbb{f}_Z$ bereits vollständig bestimmt.

An Hand von Beispielen werden wir nun zeigen, wie sich die Verteilungsdichte $\mathbb{f}_Z$ einer diskreten Zufallsvariablen Z ermitteln lässt. Es werden sich dabei auch Verteilungsdichten ergeben, welche nicht in *Mathematica* implementiert sind. Wir werden zeigen, wie sich diese Verteilungsdichten unter Verwendung von *Mathematica* behandeln lassen.

13.3.2 Beispiel: In Urne I befinden sich 3 rote und 7 schwarze Kugeln. In Urne II befindet sich 7 rote und 3 schwarze Kugeln. Aus diesen Urnen wird nun obwechselnd (beginnend mit Urne I) so lange jeweils eine Kugel gezogen, bis erstmals eine rote Kugel gezogen wird. Beobachtet wird die Anzahl X (bzw Y) der dabei aus der ersten (bzw zweiten) Urne gezogenen Kugeln. Man ermittle die Verteilungsdichte $\mathbb{f}_X$ (bzw $\mathbb{f}_Y$) und stelle die Ergebnisse als Histogramm dar, für den Fall,
a) dass die Kugeln jeweils zurück gelegt werden,
b) dass die Kugeln nicht zurück gelegt werden.

▼

Lösung: i) Wir bezeichnen mit

A_i das Ereignis “beim i-ten Zug wird aus Urne I zum ersten Mal eine rote Kugel gezogen” und mit

B_i das Ereignis “beim i-ten Zug wird aus Urne II zum ersten Mal eine rote Kugel gezogen”.

Die Ereignisse A_1, A_2,... und B_1, B_2,... sind offenbar vollständig unabhängig.

a) Es wird angennomen, daß die Kugeln jeweils zurück gelegt werden, so folgt

$$\mathbb{P}[A_i] = \mathbb{P}[A] = \frac{3}{10}, \mathbb{P}[B_i] = \mathbb{P}[B] = \frac{7}{10}, \; i = 1, 2, 3, \ldots$$

ii) Die Zufallsvariable X kann die Werte $x \in \mathbb{T}_X := \{1, 2, 3, \ldots\}$ annehmen und ist damit *diskret*.

Für die Menge $\{X = k\}$ erhält man also

$$\mathbb{P}[\{X = 1\}] = \mathbb{P}[A \cup B] = \mathbb{P}[A] + \mathbb{P}[B] - \mathbb{P}[A]\,\mathbb{P}[B]$$

$$\mathbb{P}[\{X = 2\}] = \mathbb{P}[A^c \cap B^c \cap (A \cup B)] = (1 - \mathbb{P}[A])\,(1 - \mathbb{P}[B])\,(\mathbb{P}[A] + \mathbb{P}[B] - \mathbb{P}[A]\,\mathbb{P}[B]$$

...

$$\mathbb{P}[\{X = k\}] = \mathbb{P}[\bigcap_{i=0}^{k-1} (A^c \cap B^c)^i \cap (A \cup B)] = [(1 - \mathbb{P}[A])\,(1 - \mathbb{P}[B])]^{k-1}\,(\mathbb{P}[A] + \mathbb{P}[B] - \mathbb{P}[A]\,\mathbb{P}[B])$$

Also gilt für die Verteilungsdichte $\mathbb{f}_X$ von X

$$\mathbb{f}_X[x] = \mathbb{P}_X[\{X = x\}] = \begin{cases} [(1 - \mathbb{P}[A])\,(1 - \mathbb{P}[B])]^{x-1}\,(\mathbb{P}[A] + \mathbb{P}[B] - \mathbb{P}[A]\,\mathbb{P}[B]) & \text{für } x \in \mathbb{T}_X := \{1, 2, 3, \ldots\}, \\ 0 & \text{sonst.} \end{cases}$$

Die Zufallsvariable Y kann die Werte $y \in \mathbb{T}_Y := \{0, 1, 2, \ldots\}$ annehmen und ist damit auch *diskret*.

Für die Menge $\{Y = k\}$ erhält man also

$$\mathbb{P}[\{Y = 0\}] = \mathbb{P}[A]$$

$\mathbb{P}[\{Y=1\}] = \mathbb{P}[(A^c \cap B) \cup (A^c \cap B^c \cap A)] = (1-\mathbb{P}[A])\,(\mathbb{P}[A] + \mathbb{P}[B] - \mathbb{P}[A]\,\mathbb{P}[B]));$

...

$\mathbb{P}[\{Y=k\}] = \mathbb{P}[\bigcap_{i=0}^{k-1} (A^c \cap B^c)^i \cap (A^c) \cap (A \cup B)] = (1 - \mathbb{P}[A])^k\,(1 - \mathbb{P}[B])^{k-1}\,(\mathbb{P}[A] + \mathbb{P}[B] - \mathbb{P}[A]\,\mathbb{P}[B])$

Also gilt für die Verteilungsdichte g_Y von Y

$$g_Y[y] = \mathbb{P}_Y[\{Y=y\}] = \begin{cases} (1-\mathbb{P}[A])^y\,(1-\mathbb{P}[B])^{y-1}\,(\mathbb{P}[A]+\mathbb{P}[B]-\mathbb{P}[A]\,\mathbb{P}[B]) & \text{für } y \in \mathbb{T}_Y := \{0,\ 1,\ 2,\ \ldots\}, \\ 0 & \text{sonst.} \end{cases}$$

iii) Die Verteilungsdichte lässt sich in *Mathematica* eingeben

```
pa[n_, k_] := k/n; pb[n_, k_] := (n - k)/n;
f[x_, n_, k_] := 0 /; Or[x < 1, NumberQ[x] && Not[IntegerQ[x]]]
f[x_, n_, k_] := ((1 - pa[n, k]) (1 - pb[n, k]))^(x-1) (pa[n, k] + pb[n, k] - pa[n, k] pb[n, k]);
TableForm[Table[{f[x, 10, 3], x}, {x, 1, 6}], TableHeadings -> {None, {"Wahrscheinlichkeit", "X"}}] //
  N
```

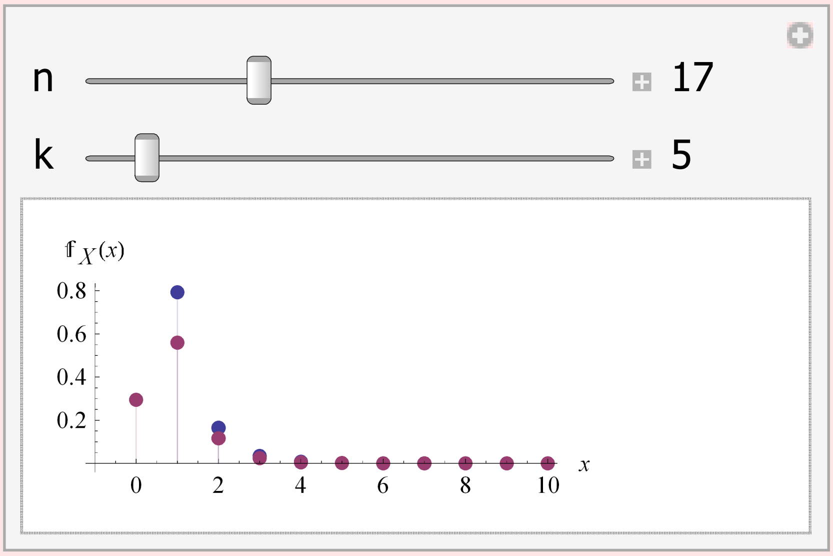
Wahrscheinlichkeit	X
0.79	1.
0.1659	2.
0.034839	3.
0.00731619	4.
0.0015364	5.
0.000322644	6.

```
pa[n_, k_] := k/n; pb[n_, k_] := (n - k)/n;
g[y_, n_, k_] := 0 /; Or[NumberQ[y] && Not[IntegerQ[y]]]
g[y_, n_, k_] := (1 - pa[n, k])^y (1 - pb[n, k])^(y-1) (pa[n, k] + pb[n, k] - pa[n, k] pb[n, k]);
g[0, n_, k_] := pa[n, k];
TableForm[Table[{g[y, 10, 3], y}, {y, 0, 6}], TableHeadings -> {None, {"Wahrscheinlichkeit", "Y"}}] //
  N
```

Wahrscheinlichkeit	Y
0.3	0.
0.553	1.
0.11613	2.
0.0243873	3.
0.00512133	4.
0.00107548	5.
0.000225851	6.

und für beliebige Wert von n und k auch graphisch durch ein Balkendiagramm darstellen:

```
Manipulate[If[k ≤ n, ListPlot[{Table[{x, f[x, n, k]}, {x, 1, 10}], Table[{y, g[y, n, k]}, {y, 0, 10}]},
    PlotStyle → PointSize[0.03], Filling → Axis, AspectRatio → 0.4, AxesOrigin → {-1, 0}, PlotRange → All,
    AxesLabel → {x, f_X[x]}, ImageSize → {200, 100}], "k>n"],
  {n, 2, 50, 1, Appearance → "Labeled"}, {k, 1, 50, 1, Appearance → "Labeled"}]
```



iv) Wir können auch überprüfen, ob diese Funktion f_Z für beliebige Werte von n tatsächlich eine diskrete Verteilungsdichte ist, ob also stets die Beziehungen

$$\sum_{x=1}^{n} f_X[x] = 1 \text{ und } \sum_{y=0}^{n} g_Y[y] = 1$$

gilt.

```
Sum[f[x, 10, 3], {x, 1, ∞}]
g[0, 10, 3] + Sum[g[y, 10, 3], {y, 1, ∞}]
```

```
1
```

```
1
```

b) Es wird angenommen, daß die Kugeln nicht zurück gelegt werden.

ii) Die Zufallsvariable X kann die Werte $x \in \mathbb{T}_X := \{1, 2, 3, 4\}$ annehmen und ist damit *diskret*.

Für die Menge $\{X = 1\}$ erhält man also

$$\mathbb{P}[\{X = 1\}] = \mathbb{P}[A_1 \cup B_1] = \mathbb{P}[A_1] + \mathbb{P}[B_1] - \mathbb{P}[A_1]\,\mathbb{P}[B_1]$$

$$\mathbb{P}[\{X = 2\}] = \mathbb{P}\left[A_1^c \cap B_1^c \cap (A_2 \cup B_2)\right] = (1 - \mathbb{P}[A_1])\,(1 - \mathbb{P}[B_1])\,(\mathbb{P}[A_2] + \mathbb{P}[B_2] - \mathbb{P}[A_2]\,\mathbb{P}[B_2])$$

...

$$\mathbb{P}[\{X = k\}] = \mathbb{P}\left[\bigcap_{i=1}^{k-1} \left(A_i^c \cap B_i^c\right) \cap \left(A_k \cup B_k\right)\right] = \prod_{i=1}^{k-1} (1 - \mathbb{P}[A_i])\,(1 - \mathbb{P}[B_i])\,\left(\mathbb{P}\left[A_k\right] + \mathbb{P}\left[B_k\right] - \mathbb{P}\left[A_k\right]\mathbb{P}\left[B_k\right]\right),$$

wobei $\mathbb{P}\left[A_k\right] = \frac{3}{11-k}$ und $\mathbb{P}\left[B_k\right] = \frac{7}{11-k}$

Also gilt für die Verteilungsdichte f_X von X

$$f_X[x] = \begin{cases} \prod_{i=1}^{k-1} (1 - \mathbb{P}[A_i])\,(1 - \mathbb{P}[B_i])\,(\mathbb{P}[A_x] + \mathbb{P}[B_x] - \mathbb{P}[A_x]\,\mathbb{P}[B_x]), & \text{für } x \in \mathbb{T}_X := \{1, 2, 3, 4\}, \\ 0 & \text{sonst.} \end{cases}$$

Die Zufallsvariable Y kann die Werte $y \in \mathbb{T}_Y := \{0, 1, 2, 3, 4\}$ annehmen und ist damit *diskret*.

Für die Menge $\{Y = 0\}$ erhält man also

$$\mathbb{P}[\{Y=0\}] = \mathbb{P}[A_1]$$

$$\mathbb{P}[\{Y=1\}] = \mathbb{P}\left[\left(A_1^c \cap B_1\right) \cup \left(A_1^c \cap B_1^c \cap A_2\right)\right] = (1-\mathbb{P}[A_1])\,(\mathbb{P}[A_2] + \mathbb{P}[B_1] - \mathbb{P}[A_2]\,\mathbb{P}[B_1])$$

...

$$\mathbb{P}[\{Y=k\}] = \mathbb{P}\left[\cap_{i=1}^{k-1}\left(A_i^c \cap B_i^c\right) \cap A_k^c \cap \left(A_{k+1} \cup B_k\right)\right] =$$

$$\prod_{i=1}^{k-1} (1-\mathbb{P}[A_i])\,(1-\mathbb{P}[B_i])\left(1-\mathbb{P}\left[A_k\right]\right)\left(\mathbb{P}\left[A_{k+1}\right] + \mathbb{P}\left[B_k\right] - \mathbb{P}\left[A_{k+1}\right]\mathbb{P}\left[B_k\right]\right)$$

Also gilt für die Verteilungsdichte g_Y von Y

$$g_Y[y] = \begin{cases} \prod_{i=1}^{k-1} (1-\mathbb{P}[A_i])\,(1-\mathbb{P}[B_i])\left(1-\mathbb{P}\left[A_k\right]\right)\left(\mathbb{P}\left[A_{k+1}\right] + \mathbb{P}\left[B_k\right] - \mathbb{P}\left[A_{k+1}\right]\mathbb{P}\left[B_k\right]\right), & \text{für } y \in \mathbb{T}_Y := \{0, 1, 2, 3, 4\}, \\ 0 & \text{sonst,} \end{cases}$$

wobei $\mathbb{P}\left[A_k\right] = \frac{3}{11-k}$ und $\mathbb{P}\left[B_k\right] = \frac{7}{11-k}$.

iii) Die Verteilungsdichte lässt sich in *Mathematica* eingeben

```
pa[j_, n_, k_] := k/(n + 1 − j);
pb[j_, n_, k_] := (n − k)/(n + 1 − j);
f[x_, n_, k_] := 0 /; Or[x > n, NumberQ[x] && Not[IntegerQ[x]]]
f[x_, n_, k_] := Product[(1 − pa[i, n, k]) (1 − pb[i, n, k]), {i, 1, x − 1}] ∗
    (pa[x, n, k] + pb[x, n, k] − pa[x, n, k] pb[x, n, k]);
tf[n_, k_] := TableForm[Table[{f[x, n, k], x}, {x, 1, k + 1}],
    TableHeadings −> {None, {"Wahrscheinlichkeit", "X"}}];
tf[10, 3] // N
```

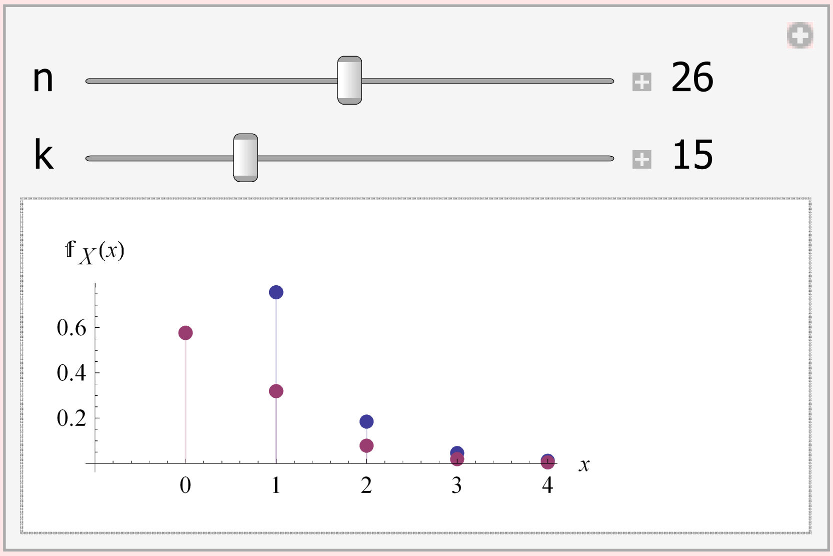
Wahrscheinlichkeit	X
0.79	1.
0.178889	2.
0.0286806	3.
0.00243056	4.

```
pa[j_, n_, k_] := k/(n + 1 − j);
pb[j_, n_, k_] := (n − k)/(n + 1 − j);
g[y_, n_, k_] := 0 /; Or[y > n − 1, NumberQ[y] && Not[IntegerQ[y]]]
g[y_, n_, k_] := Product[(1 − pa[i, n, k]) (1 − pb[i, n, k]), {i, 1, y − 1}] ∗ (1 − pa[y, n, k]) ∗
    (pa[y + 1, n, k] + pb[y, n, k] − pa[y + 1, n, k] pb[y, n, k]);
g[0, n_, k_] = pa[1, n, k];
tg[n_, k_] := TableForm[Table[{g[y, n, k], y}, {y, 0, k + 1}],
    TableHeadings −> {None, {"Wahrscheinlichkeit", "Y"}}];
tg[10, 3] // N
```

Wahrscheinlichkeit	Y
0.3	0.
0.56	1.
0.120556	2.
0.0180556	3.
0.00138889	4.

und für beliebige Wert von n und k auch graphisch durch ein Balkendiagramm darstellen:

```
Manipulate[If[k ≤ n, ListPlot[{Table[{x, f[x, n, k]}, {x, 1, 4}], Table[{y, g[y, n, k]}, {y, 0, 4}]},
    PlotStyle → PointSize[0.03], Filling → Axis, AspectRatio → 0.4, AxesOrigin → {−1, 0}, PlotRange → All,
    AxesLabel → {x, 𝕗X[x]}, ImageSize → {200, 100}], "k>n"],
  {n, 2, 50, 1, Appearance → "Labeled"}, {k, 1, 50, 1, Appearance → "Labeled"}]
```



iv) Wir können auch überprüfen, ob diese Funktione $\mathbb{f}_X$ und $\mathbb{g}_Y$ für beliebige Wert von n tatsächlich diskrete Verteilungsdichte sind, ob also stets die Beziehungen $\sum_{x=1}^{n} \mathbb{f}_X[x] = 1$ und $\sum_{x=1}^{n} \mathbb{g}_Y[y] = 1$ gelten.

```
Sum[f[x, 10, 3], {x, 1, 4}]
g[0, 10, 3] + Sum[g[y, 10, 3], {y, 1, 4}]
```

```
1
```

```
1
```

13.3.3 Beispiel: Unser Zufallsexperiment besteht im Werfen von zwei homogenen Würfeln. Mit der Zufallsvariablen Z wird die dabei auftretenden Augensumme bezeichnet. Man analysiere die Verteilungsdichte $\mathbb{f}_Z$ von Z.

▼

Lösung:

a) Als Ereignisraum für unser Zufallsexperiment bietet sich die Menge (vgl dazu Beispiel 3.2.2 und Besipiel 12.2.3)

$$\Omega = \{\{x_1, x_2\} \mid x_i \in \{1, 2, 3, 4, 5, 6\}\}$$

an, wobei alle Realisierungen gleichwahrscheinlich sind.

b) Bei der Zufallsvariablen Z handelt es sich um die Abbildung

$$Z : \Omega \to \mathbb{R} \quad \text{mit} \quad Z[\{x_1, x_2\}] = x_1 + x_2$$

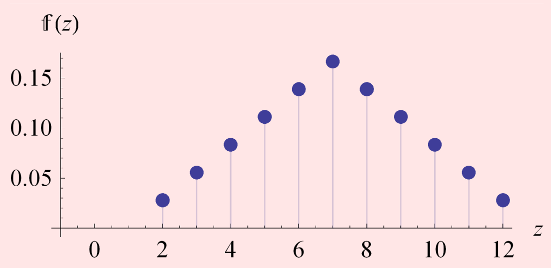
Die Zufallsvariable Z kann offenbar nur die Werte $z \in \mathbb{T}_Z = \{2, 3, \ldots, 12\}$ annehmen uns ist damit *diskret*. Für alle $z \in \mathbb{T}_Z$ entspricht das Ereignis $\{Z = z\}$ der Menge

$$\{Z = z\} = \{\{x_1, x_2\} \in \Omega \mid x_1 + x_2 = z\}$$

Also gilt für die Verteilungsdichte $\mathbb{f}_Z$ von Z

$$\mathbb{f}_Z[z] = \mathbb{P}[\{Z = z\}] = \begin{cases} (z-1)/6^2 & \text{für } z \in \{2, 3, \ldots, 7\} \\ (13-z)/6^2 & \text{für } z \in \{8, 9, \ldots, 12\} \\ 0 & \text{sonst} \end{cases}$$

```
fd1[z_] := 0 /; Or[z < 2, z > 12, NumberQ[z] && Not[IntegerQ[z]]]
fd1[z_] := (z - 1) / 6^2 /; 2 ≤ z ≤ 7
fd1[z_] := (13 - z) / 6^2 /; 8 ≤ z ≤ 12

ListPlot[Table[{z, fd1[z]}, {z, 2, 12}], PlotStyle → PointSize[0.03], Filling → Axis, AspectRatio → 0.4,
  AxesOrigin → {-1, 0}, PlotRange → All,
  AxesLabel → {z, f[z]}, ImageSize → {200, 100}]
```



13.3.4 Beispiel: Durch einen Kanal werden n unabhängige Kodeworte übertragen (vgl. Beispiel 11.2.5). Die Wahrscheinlichkeit dafür, dass das Kodewort richtig empfangen wird, ist 0.9. Mit der Zufallsvariablen Z wird die Anzahl der Kodeworten, die richtig empfangen wurden, bezeichnet. Man analysiere die Verteilungsdichte $\mathbb{f}_Z$ von Z.

▼

Lösung:

Die Zufallsvariable Z kann offenbar nur die Werte $z \in \mathbb{T}_Z = \{1, 2, \ldots, n\}$ annehmen uns ist damit diskret. Aus dem Beispiel

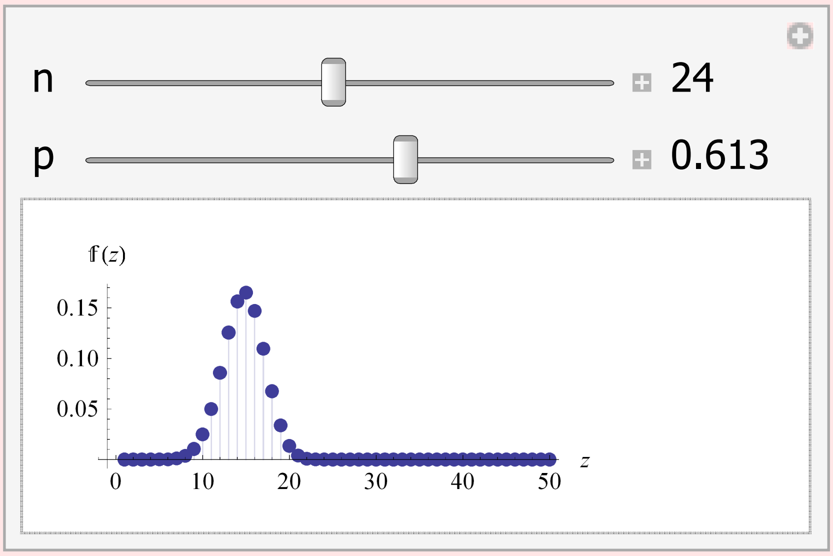
$$\mathbb{f}_Z[z] = \mathbb{P}[\{Z = z\}] = \begin{cases} \binom{n}{z} 0.9^z\, 0.1^{n-z} & \text{für } z \in \mathbb{T}_Z \\ 0 & \text{sonst} \end{cases}$$

Diese Verteilungsdichte lässt sich mühelos in *Mathematica* eingeben

```
fd1[z_, n_, p_] := 0 /; Or[z < 1, z > n, NumberQ[z] && Not[IntegerQ[z]]]
fd1[z_, n_, p_] := Binomial[n, z] p^z (1 - p)^(n - z)
```

und mit Hilfe von Manipulate für beliebige Werte von p und n auf dynamische Weise graphisch darstellen:

```
Manipulate[ListPlot[Table[{z, fd1[z, n, p]}, {z, 1, 50}],
   PlotStyle → PointSize[0.03], Filling → Axis, AspectRatio → 0.4, AxesOrigin → {−1, 0}, PlotRange → All,
   AxesLabel → {z, 𝕗[z]}, ImageSize → {200, 100}],
 {n, 1, 50, 1, Appearance → "Labeled"}, {p, 0, 0.99, Appearance → "Labeled"}]
```



Gelegentlich bereitet die theoretische Berechnung der Verteilungsdichte einer diskreten Zufallsvariablen Probleme. In diesen Fällen kann es sinnvoll sein, diese Verteilungsdichte durch Simulation näherungsweise zu ermitteln. Man erzeugt dazu eine genügend große Anzahl (vergleiche dazu unsere Faustregel) von Realisierungen dieser Zufallsvariablen Z und ermittelt von diesem Datenmaterial daten unter Verwendung des Befehls DiscreteEmpiricalPDF die diskrete empirische Verteilungsdichte. Für alle $z \in \mathbb{R}$ stellt die diskrete empirische Verteilungsdichte

$$\hat{\mathbb{f}}_Z[z] = \text{relative Häufigkeit des Ereignisses } \{z\} \text{ im Datenmaterial daten}$$

eine gute Näherung für die gesuchte Verteilungsdichte der diskreten Zufallsvariablen Z dar.

■ DiscreteEmpiricalPDF[*daten*, *z*]

ordnet jedem $z \in \mathbb{R}$ die relative Häufigkeit des Ereignisses $\{z\}$ im Datenmaterial *daten* zu.

13.3.5 Beispiel: Aus einer Urne mit $s = 8$ schwarzen, $r = 6$ roten und $g = 4$ grünen Kugeln werden so lange Kugeln gezogen und nach jedem Zug wieder zurückgelegt, bis erstmals hintereinander zwei gleich gefärbte Kugeln gezogen werden (vergleiche dazu Beispiel 4.2.18). Die Zufallsvariable Z gibt an, wieviele Züge dazu notwendig sind. Man bestimme die Verteilungsdichte $\mathbb{f}_Z$ von Z und ermittle damit die Wahrscheinlichkeit dafür, dass maximal $m = 5$ Züge notwendig sind, bis erstmals hintereinander zwei gleich gefärbte Kugeln gezogen werden.

▼

Lösung: a) Die Zufallsvariable Z kann nur Werte $z \in \mathbb{T}_Z = \{2, 3, \ldots\}$ annehmen und ist damit diskret. Die exakte Berechnung der Wahrscheinlichkeiten $\mathbb{f}_Z[z] = \mathbb{P}[\{Z = z\}]$ für $z \in \mathbb{T}_Z$ ist kompliziert. Wir wollen deshalb diese Wahrscheinlichkeiten näherungsweise durch Simulation ermitteln, indem wir (analog zu Beispiel 4.2.18) viele mögliche Realisierungen dieser Zufallsvariablen Z erzeugen und von diesem Datenmaterial anschließend die relativen Häufigkeiten aller Werte $z \in \mathbb{T}_Z$ ermitteln.

b) Wir simulieren dazu zuerst die Farbe farbe der gezogenen Kugel, ermitteln damit $n = 10^4$ mögliche Zugserien, bestimmen für jede dieser Zugserien die Anzahl der Züge, welche notwendig sind, bis erstmals zweimal hintereinander Kugeln der gleichen Farbe gezogen werden und nennen dieses Datenmaterial daten:

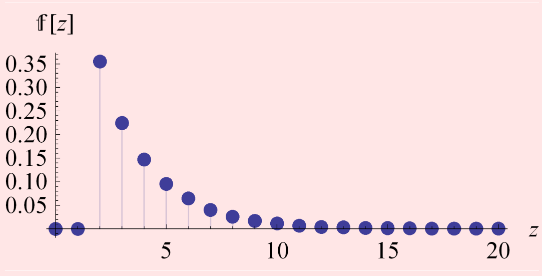
```
s = 8; r = 6; g = 4; n = 10^4;
farbe[s_, r_, g_] := Which[x = RandomInteger[{1, s + r + g}]; x ≤ s, schwarz, x ≤ s + r, rot, x > s + r, grün]
daten = Table[For[zug = Table[farbe[s, r, g], {2}], zug[[−1]] =!= zug[[−2]], zug = Append[zug, farbe[s, r, g]]];
    Length[zug], {n}];
```

c) Mit dem Befehl DiscreteEmpiricalPDF ermitteln wir nun die diskrete empirische Verteilungsdichte der Zufallsvariablen Z (dabei ist zu beachten, dass diese diskrete empirische Verteilungsdichte vom Datenmaterial daten abhängt und sich mit jedem neuen Aufruf von daten daher geringfügig ändern wird)

```
fd4[z_] := DiscreteEmpiricalPDF[daten, z]
```

und zeichnen diese diskrete empirische Verteilungsdichte:

```
ListPlot[Table[{z, fd4[z]}, {z, 0, 20}],
  PlotStyle → PointSize[0.03], Filling → Axis, AspectRatio → 0.4, AxesOrigin → {0, 0}, PlotRange → All,
  AxesLabel → {z, 𝕗[z]}, ImageSize → {200, 100}]
```



d) Für die Wahrscheinlichkeit dafür, dass maximal $m = 5$ Züge notwendig sind, bis erstmals hintereinander zwei gleich gefärbte Kugeln gezogen werden, gilt somit (näherungsweise)

```
m = 5;
Sum[fd4[z], {z, 2, m}] // N
```

```
0.8218
```

```
Clear[s, r, g, n, x, zug, farbe, daten, m, fd4]
```