Erzeugende Funktion der Binomialverteilung

Erzeugende Funktion der geometrischen Verteilung

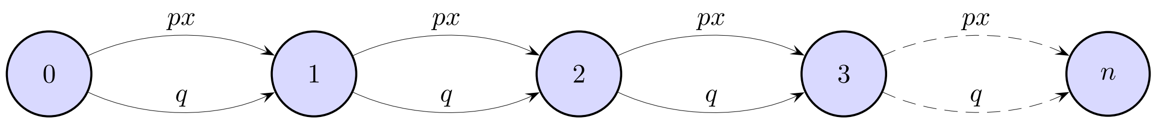
Warten auf den n-ten Erfolg

Markierungsmethode

Reduktion von Graphen

Wartezeit

↑ Erzeugende Funktion der Binomialverteilung



Die Bernoulli-Kette habe die Länge $n = 8$.
Jedem Pfad wird das Produkt der Pfadbezeichnungen wie z. B. $p^6q^2x^6$ zugeordnet. Die erzeugende Funktion $g(x)$ ist die Summe aller dieser Terme. In der nach Potenzen von x geordneten Form werden die Wahrscheinlichkeiten der Pfade mit jeweils gleicher Trefferzahl k durch den Koeffizienten p_k von x^k zusammengefasst.

$$g(x) = (px+q)^n = p^8x^8 + 8p^7qx^7 + 28p^6q^2x^6 + 56p^5q^3x^5 + 70p^4q^4x^4 + 56p^3q^5x^3 + 28p^2q^6x^2 + 8pq^7x + q^8$$

$$g(x) = (px+q)^n = p_8x^8 + p_7x^7 + p_6x^6 + p_5x^5 + p_4x^4 + p_3x^3 + p_2x^2 + p_1x^1 + p_0x^0$$

Erwartungswert und Varianz einer Verteilung lassen sich mit der geschlossenen Darstellung einer erzeugenden Funktion leicht ermitteln.

$$g'(x) = 8p_8x^7 + 7p_7x^6 + 6p_6x^5 + 5p_5x^4 + 4p_4x^3 + 3p_3x^2 + 2p_2x^1 + 1p_1x^0$$

$$\Longrightarrow \quad g'(1) = 8p_8 + 7p_7 + 6p_6 + 5p_5 + 4p_4 + 3p_3 + 2p_2 + 1p_1 = E(X) = \mu$$

$$g''(x) = 7\cdot 8p_8x^6 + 6\cdot 7p_7x^5 + 5\cdot 6p_6x^4 + 4\cdot 5p_5x^3 + 3\cdot 4p_4x^2 + 2\cdot 3p_3x^1 + 1\cdot 2p_2x^0$$

$$\Longrightarrow \quad g''(1) = 7\cdot 8p_8 + 6\cdot 7p_7 + 5\cdot 6p_6 + 4\cdot 5p_5 + 3\cdot 4p_4 + 2\cdot 3p_3 + 1\cdot 2p_2$$

$$\Longrightarrow \quad g'(1) + g''(1) = E(X^2)$$

$$\Longrightarrow \quad V(X) = E(X^2) - \mu^2 = g'(1) + g''(1) - (g'(1))^2$$

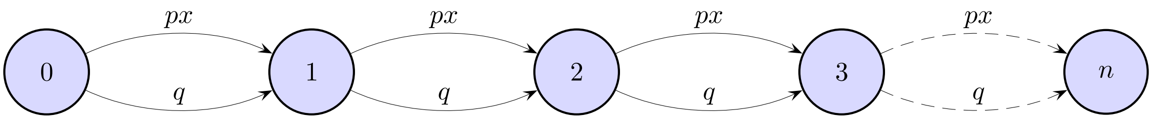
Für die Binomialverteilung gilt dann:

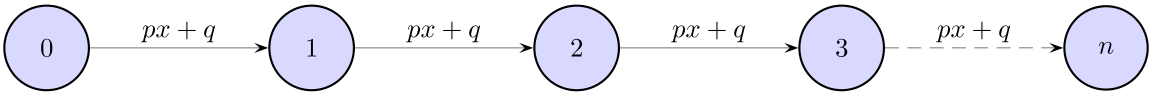
$$g'(x) = np(px+q)^{n-1}$$
$$g'(1) = np = \mu$$

$$g''(x) = n(n-1)p^2(px+q)^{n-2}$$
$$g''(1) = n(n-1)p^2$$

$$\Longrightarrow \quad V(X) = g'(1) + g''(1) - (g'(1))^2 = np + n(n-1)p^2 - n^2p^2 = np - np^2 = npq$$

↑ Erzeugende Funktion der Binomialverteilung



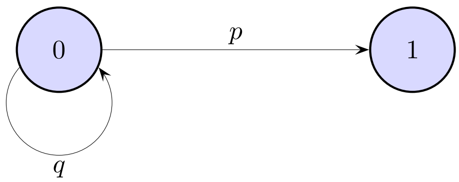


$$g(x) = (px+q)^n = p_8x^8 + p_7x^7 + p_6x^6 + p_5x^5 + p_4x^4 + p_3x^3 + p_2x^2 + p_1x^1 + p_0x^0, \qquad n = 8$$

Um für komplexere Graphen die erzeugende Funktion zu ermitteln, können parallele, gleichgerichtete Zweige zusammengelegt werden.

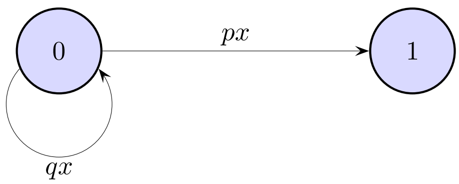
Man wird allgemein die Wahrscheinlichkeiten derjenigen Übergänge mit einem x multiplizieren, wenn von Interesse ist, mit welcher Wahrscheinlichkeit der Übergang $k$mal durchlaufen wird. Gibt die Zufallsvariable X die Anzahl der Durchläufe an, so beinhaltet die erzeugende Funktion die Verteilung $p_k = P(X = k)$ von X.

↑ Erzeugende Funktion der geometrischen Verteilung



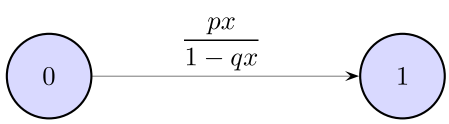
Ein Glücksrad wird solange gedreht bis zum 1. Mal eine 1 erscheint. X sei die Anzahl der Übergänge bis zur Absorption. Dem Graphen kann die Verteilung $p_k = q^{k-1}p$ unmittelbar abgelesen werden.

Die für komplexere Graphen erforderliche Schleifenbeseitigung kann an diesem einfachen Beispiel erläutert werden. Da alle Übergänge gezählt werden sollen, sind die Wahrscheinlichkeiten beider Übergänge mit einem x zu multiplizieren.



Als erzeugende Funktion erhalten wir dann
(beachte die Summenformel für eine geometrische Reihe):

$$g(x) = px + pqx^2 + pq^2x^3 + pq^3x^4 + \ldots = \frac{px}{1-qx}$$



Mit

$$g'(1) = \frac{p}{(1-q)^2}$$

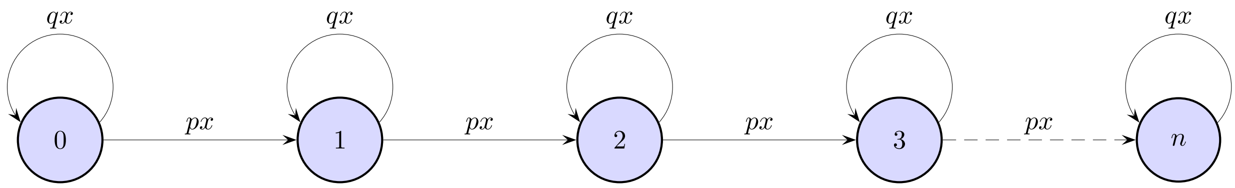
$$g''(1) = \frac{2pq}{(1-q)^3}$$

erhalten wir:

$$\mu = \frac{1}{p}$$

$$\sigma^2 = \frac{q}{p^2}$$

↑ Warten auf den n-ten Erfolg

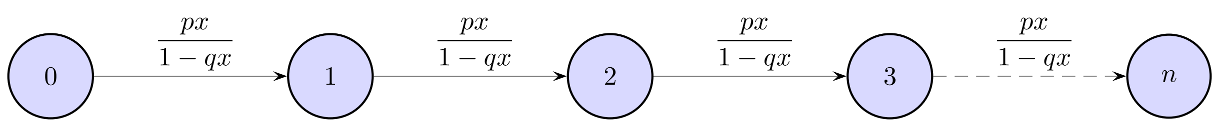


Ein Glücksrad wird solange gedreht bis zum n-ten Mal eine 1 erscheint.
X sei die Anzahl der Übergänge bis zur Absorption.

Dieses Problem ist vom Sammlerproblem zu unterscheiden. Aus einer Urne mit n nummerierten Kugeln wird zufällig jeweils eine gezogen und wieder zurückgelegt. Dies wird solange wiederholt, bis jede Kugel mindestens einmal gezogen wurde. Hierbei verringern sich die Übergangswahrscheinlichkeiten

$$p_i = \frac{n+1-i}{n}, \;\; i = 1, 2, 3, \ldots, n \qquad \text{gemäß der abnehmenden Anzahl noch nicht gezogener Kugeln.}$$

Für das n-te-Erfolgs-Problem ist die mittlere Wartezeit (Schrittzahl) $E(X)$ bis zur Absorption in n gesucht.
Wir beseitigen die Schleifen:



Die erzeugende Funktion für X lautet dann:

$$g(x) = \Big(\frac{px}{1-qx}\Big)^n$$

Mit

$$g'(1) = \frac{n}{p}$$

$$g''(1) = \frac{n(n+1-2p)}{p^2}$$

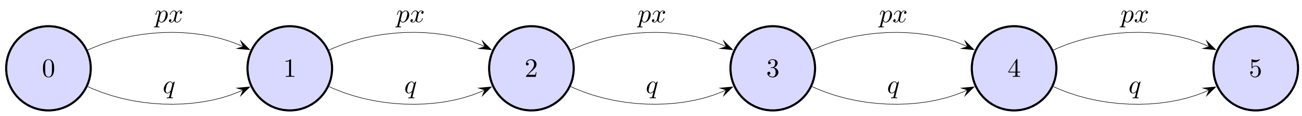
erhalten wir:

$$\mu = \frac{n}{p}$$

$$\sigma^2 = \frac{nq}{p^2}$$

Dieses Ergebnis ist mit der Zerlegung $X = X_1 + X_2 + \ldots + X_n$ unmittelbar einsichtig.

↑ Wahrscheinlichkeits-Interpretation der erzeugenden Funktion

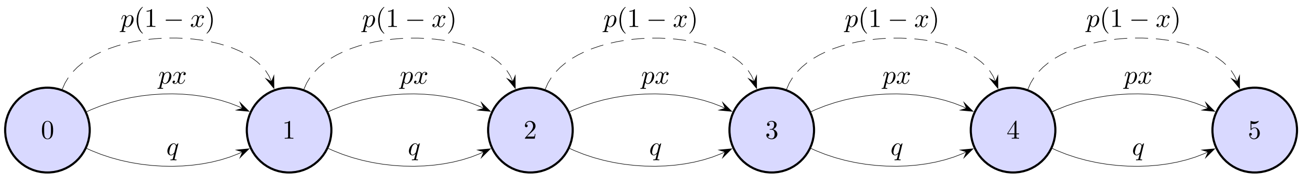


Die Bernoulli-Kette habe die Länge $n = 5$.

Um die erzeugende Funktion $g(x)$ zu ermitteln, wird jedem Pfad das Produkt der Pfadbezeichnungen längs des Pfades wie z.B. $px \cdot px \cdot q \cdot px \cdot q = p^3q^2x^3$ zugeordnet und die Summe aller dieser Terme gebildet. Wenn wir die Pfadbezeichnungen als Wahrscheinlichkeiten ansehen, stellt $g(x)$ die Wahrscheinlichkeit aller Pfade dar. Diese Interpretation (van Dantzig, siehe Wahrscheinlichkeitsrechnung und Statistik, Arthur Engel, Band 2, Klett Verlag 1976) ist manchmal nützlich, um einem Graphen die erzeugende Funktion ablesen zu können.

$$g(x) = (px + q)^5 = p^5x^5 + 5p^4qx^4 + 10p^3q^2x^3 + 10p^2q^3x^2 + 5pq^4x + q^5$$

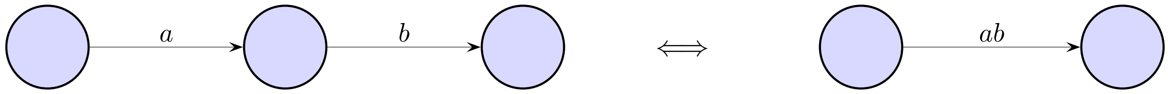
Damit nun wirklich Wahrscheinlichkeiten vorliegen können, muss der Graph - in Gedanken -
an allen px-Übergängen ergänzt werden.



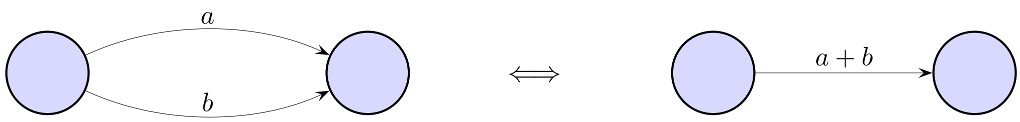
Die - gestrichelte - Ergänzung wird uns nicht weiter interessieren, sie dient lediglich dazu, der erzeugenden Funktion eine anschauliche Bedeutung zu geben.

↑ Reduktion von Graphen

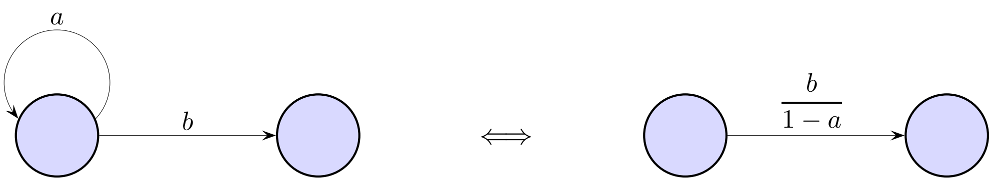
1.



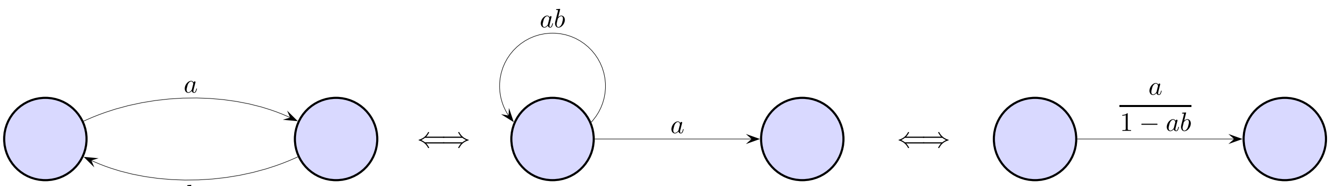
2.



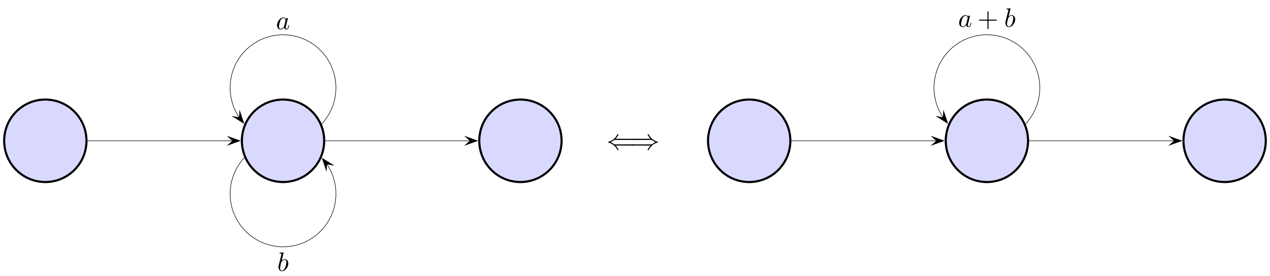
3.



4.



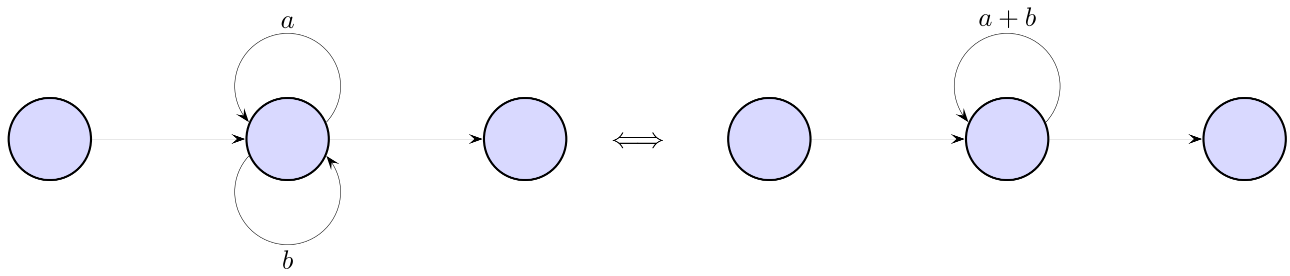
5.



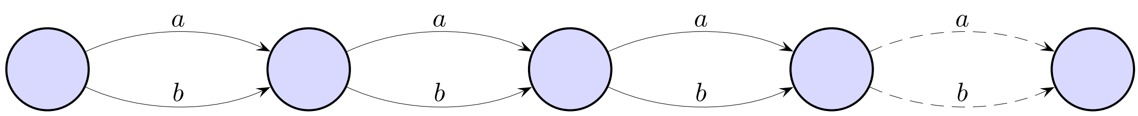
↑ Reduktion von Graphen

Mit diesen Regeln können Graphen zur Ermittlung erzeugender Funktionen und Graphen von absorbierenden Markow-Ketten vereinfacht werden. In beiden Fällen werden Summen von Produkt-Termen betrachtet, die längs der Pfade gebildet werden.

Somit sind 1. und 2. unmittelbar einsichtig.
Die Beseitigung der Schleife in 3. wurde schon bei der geometrischen Verteilung erläutert.
Der linke Graph in 4. beinhaltet eine ab-Schleife und einen a-Übergang.

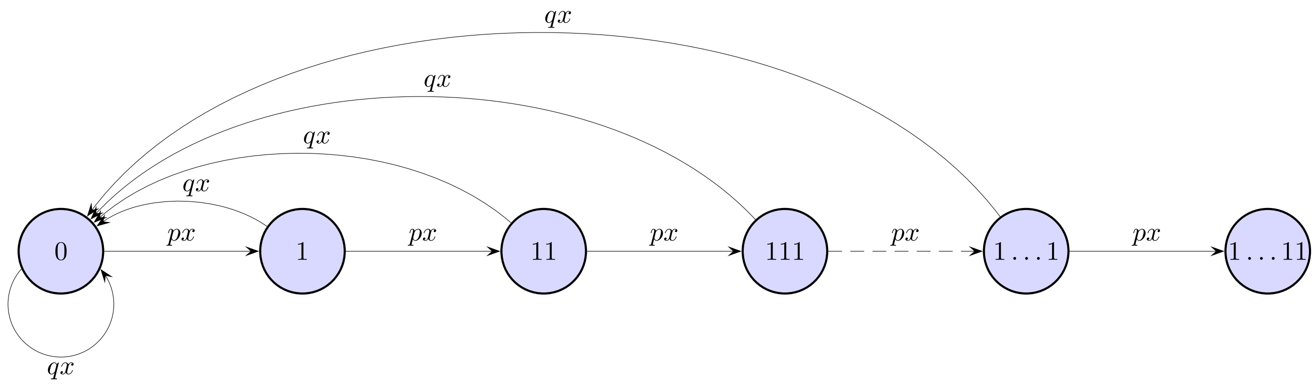


Die Doppelschleife in 5. enthält viele mögliche Pfade. Beim einmaligen Durchlaufen der Schleife entsteht die Termsumme $a+b$, beim zweimaligen Durchlaufen entsteht $(a+b)^2 = a^2 + 2ab + b^2$, beim dreimaligen Durchlaufen entsteht $(a+b)^3 = a^3 + 3a^2b + 3ab^2 + b^3$, usw.

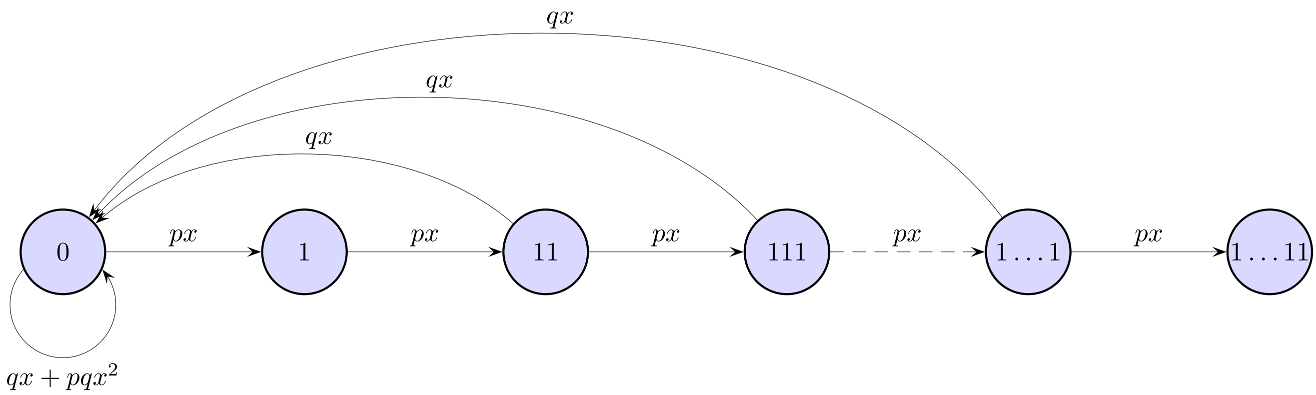


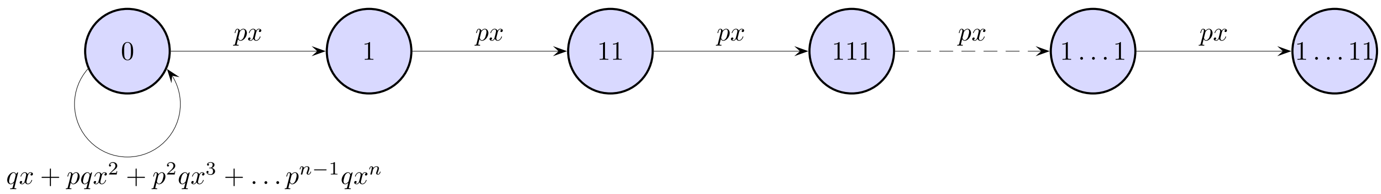
↑ Wartezeit

Die Zeichen 1 und 0 werden pro Takt mit den Wahrscheinlichkeiten p bzw. $q = 1 - p$ erzeugt.
Wir untersuchen die Wartezeit für n Treffer in Folge.



Der Graph wird reduziert.





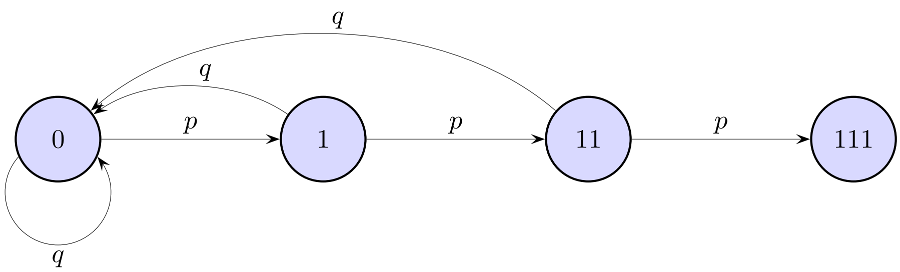
$$g(x) = \frac{p^n x^n}{1 - qx - pqx^2 - p^2qx^3 - \dots p^{n-1}qx^n} = \dots = \frac{(1-px)p^n x^n}{1 - x + qp^n x^{n+1}}$$

$$\mu = g'(1) = \frac{1-p^n}{qp^n} \sim \frac{1}{qp^n}$$

Für $n \longrightarrow \infty$ gilt $\sigma \sim \mu$, siehe Wahrscheinlichkeitsrechnung u. Statistik, A. Engel, Band 2, Klett 1976

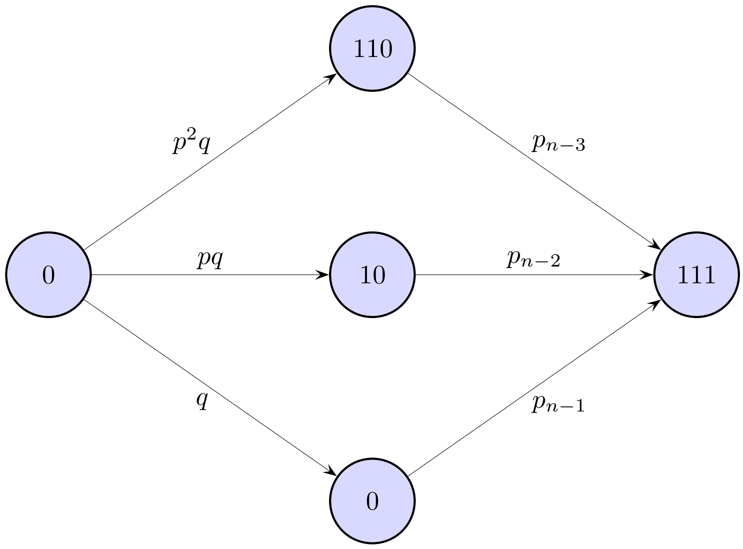
↑ Wartezeit

Die Verteilung der Wartezeit soll im einfachen Fall $n = 3$ näher untersucht werden.
Sei $p_n = P(X = n)$ die Wahrscheinlichkeit für genau n Schritte bis zur Absorption.



Für $n > 3$ muss zu Beginn eine der drei Schleifen durchlaufen werden.
Daraus ergibt sich die - für nummerische Berechnungen geeignete - rekursive Beziehung:

$$p_n = qp_{n-1} + pqp_{n-2} + p^2qp_{n-3} \qquad \text{mit } \ p_1 = p_2 = 0, \ p_3 = p^3$$



Die erzeugende Funktion lautet:

$$g(x) = \frac{(1-px)p^3x^3}{1-x+qp^3x^4} = p^3x^3 + (-p^4+p^3)x^4 + (-p^4+p^3)x^5 + (-p^4+p^3)x^6 + (-p^6q - p^4 + p^3)x^7$$
$$+ (qp^7 - 2p^6q - p^4 + p^3)x^8 + (2qp^7 - 3p^6q - p^4 + p^3)x^9 + \ldots$$

Beachtenswert: $p_4 = p_5 = p_6$

Für genauere Untersuchungen wird die erzeugende Funktion mit einer Partialbruchzerlegung umgeformt und die Summanden in geometrische Reihen entwickelt. Für Näherungen sind die für $n \longrightarrow \infty$ dominierenden Terme zu ermitteln.