

3 Galton x ZSG Movire-Laplace

Das Galtonbrett lässt sich wie folgt formal mathematisch formulieren:

Definition 3.1: Modell des Galtonbretts

Sei $(\Omega, \mathcal{F}, P)$ ein Wahrscheinlichkeitsraum.

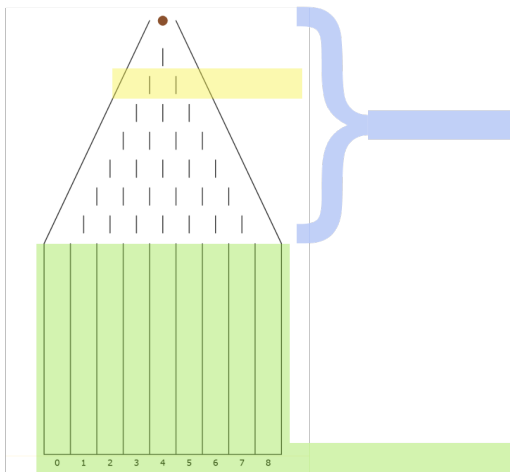
Der Fall einer Murmel durch ein Galtonbrett mit $n \in \mathbb{N}$ Reihen wird modelliert durch eine Folge von stochastisch unabhängigen und identisch verteilten Zufallsvariablen $X_1, X_2, \dots, X_n$, wobei $\mathbf{X}_i \in \{0, 1\}$. Dabei beschreibt $\mathbf{X}_i = 1$ den Fall nach rechts in der i -ten Reihe und $\mathbf{X}_i = 0$ den Fall nach links. Die Wahrscheinlichkeit sei $P(X_i = 1) = p$ und $P(X_i = 0) = 1 - p = q$. Bei einem symmetrischen Brett gilt $p = q = 0.5$.

Die Endposition der Murmel im Fach $k \in \{0, 1, \dots, n\}$ wird durch die Summe der Rechtsablenkungen beschrieben. Wir definieren die Zufallsvariable:

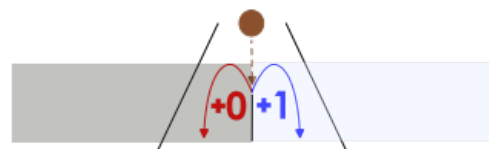
$$S_n = \sum_{i=1}^n X_i$$

Hierbei ist S_n Binomialverteilt.

Aufgabe 1: Um die Definition genauer zu verstehen, ordne die in der Definition hervorgehobenen Teile den folgenden Abbildungen des Galtonbretts zu:



Grafik 1: Darstellung des Galtonbretts nach Brünner (2025)



Grafik 2: Darstellung des Falls einer Murmel im Galtonbrett.

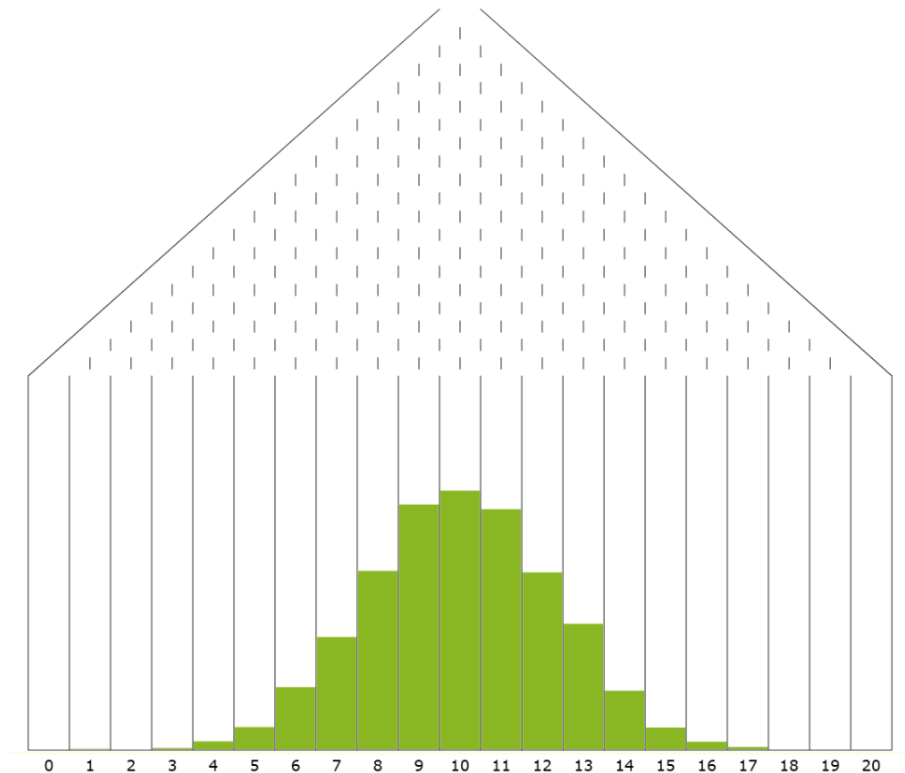
Aufgabe 2: Sobald sehr viele Murmeln durch das Galtonbrett gewandert sind, entsteht eine Verteilung, wie in Grafik 3 für ein symmetrisches Brett dargestellt wird.

1. Probiere das ganze selbst einmal auf folgender Website aus: <https://www.arndt-bruenner.de/mathe/scripts/galtonbrett.htm>.

Was fällt dir auf?

Achtung: stelle $p = 0,5$ für ein symmetrisches Brett ein!

2. Recherchiere selbstständig und zeichne den Graphen der Normalverteilung über die Verteilung des Galtonbrettes ein:



Grafik 3: Galtonbrett nach c.a. 7000 durchgefallenen Murmeln nach Brünner (2025)

3. Wiederhole oder Recherchiere selbstständig die Eigenschaften der Normalverteilung und Vergleiche diese mit der Binomialverteilung (die Verteilung des Galtonbrettes):

Normalverteilung	Binomialverteilung

Wir wollen nun das schöne mathematische Phänomen, welches sich in Aufgabe 2.2 offenbart, genauer betrachten. Hierzu formulieren wir dies zuerst mathematisch formal sauber:

Satz 3.2: Zentraler Grenzwertsatz von de Moivre-Laplace

Sei $S_n \sim \text{Bin}_{n,p}$ die Anzahl der Erfolge bei n unabhängigen Bernoulli-Versuchen mit Erfolgswahrscheinlichkeit $p \in (0, 1)$ und sei $q = 1 - p$ die Gegenwahrscheinlichkeit. (Beispielsweise bei einem Galtonbrett)

Für große n lässt sich die Zähldichte von S_n durch die Dichtefunktion der Normalverteilung annähern:

$$S_n \sim \mathcal{N}(\mu, \sigma^2)$$

wobei $\mathbb{E}(X) = \mu$, $\text{Var}(X) = \sigma^2$ und $\frac{k-np}{\sqrt{npq}}$ begrenzt ist.

Diesen Satz wollen wir nun Semi-Formal beweisen.

Aufgabe 3: Fülle die Lücken im Beweis aus. Wenn du dir unsicher bist kannst du dir alle fehlenden

Textbausteine auf dem Zusatzblatt ansehen.

1. Bevor wir den Satz beweisen, ist es wichtig, die Aussage zu verstehen. Formuliere die Kernaussage des Satzes kurz in eigenen Worten:

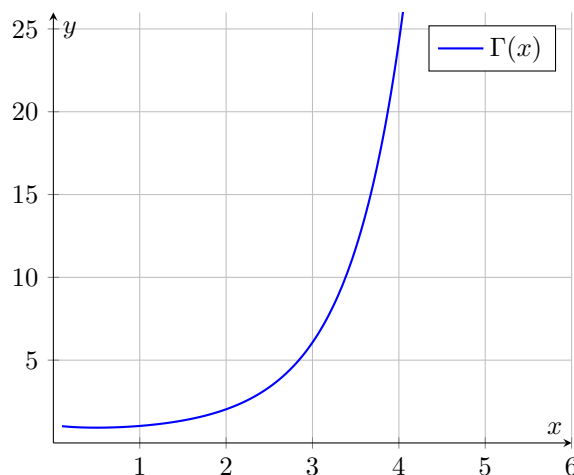
2. Zuerst muss uns nochmal klar werden, was unser Ausgangspunkt und das Ziel ist:

Ausgangspunkt: $P(S_n = k) =$

Ziel: $\frac{1}{\sqrt{2\pi\sigma^2}} \cdot e^{-\frac{(x-\mu)^2}{2\sigma^2}} =$

3. Wir wollen die Binomialverteilung als diskrete Verteilung in eine stetige Verteilung überführen. Der Aspekt der Binomialverteilung, welcher die Diskretheit forciert, ist die Fakultät. Die Fakultät der Reellen Zahlen ist die so genannte Gamma (Γ) Funktion.

Zeichne in dem folgenden Koordinatensystem, in dem auch die Gammafunktion abgebildet ist, die Funktionswerte der Fakultät (welche nur auf $\mathbb{N}$ definiert ist) für $x \in \{0, 1, 2, 3, 4, 5\}$ ein:



Mithilfe dieser Funktion lässt sich die Fakultät leicht approximieren. Diese Approximation nennt sich Sterling Approximation und besagt:

$$n! \sim \sqrt{2\pi n} \left(\frac{n}{e}\right)^n \sim \Gamma(n)$$

Setzen wir das in die Gleichung für $P(S_n = k)$ ein, erhalten wir nach umstellen:

$$P(S_n = k) \sim \frac{1}{\sqrt{2\pi}} \sqrt{\frac{n}{k(n-k)}} \left(\frac{np}{k}\right)^k \left(\frac{nq}{n-k}\right)^{n-k}$$

4. Aus dieser Formel lässt sich nun leicht schon einmal der Vorfaktor der Normalverteilung erschließen. Hierzu und zur Vereinfachung weiterer Rechnungen standardisieren wir S_n . Dies entspricht auch einer einfachen Substitution und daher vollkommen legitim. Hierzu definieren wir:

$$Z_n = \frac{S_n - E(S_n)}{\sigma(S_n)}$$

nimmt Z_n die folgenden Werte an:

$$z =$$

Nach dem Umstellen nach k und $n - k$ folgt dann:

$$k =$$

$$n - k =$$

Multiplizieren wir nun diese beiden Ausdrücke miteinander erhalten wir:

$$k(n-k) = n^2pq + (nz\sqrt{nqp}(q-p) - z^2npq)$$

Da wir uns nur für eine große Anzahl an Durchführungen also große n ($n \rightarrow \infty$) interessieren, ist beobachtbar, dass für große n der lineare Teil kaum noch einen Einfluss auf den Term hat. Daher kann dieser in diesem Fall weggelassen werden. Daher gilt:

$$k(n-k) \sim n^2 pq$$

Setzen wir dies nun in unsere diskrete Variante von $P(S_n = k)$ ein, erhalten wir den passenden Vorfaktor:

$$\begin{aligned} P(S_n = k) &\sim \frac{1}{\sqrt{2\pi}} \sqrt{\frac{n}{k(n-k)}} \left(\frac{np}{k}\right)^k \left(\frac{nq}{n-k}\right)^{n-k} \\ &\sim \text{[Empty box]} \\ &= \text{[Empty box]} \\ &= \text{[Empty box]} \\ &= \frac{1}{\sqrt{2\pi npq}} \left(\frac{k}{np}\right)^{-k} \left(\frac{n-k}{nq}\right)^{-n+k} \end{aligned}$$

5. Zum Schluss müssen wir noch die e -Funktion mit in die Gleichung bekommen. Das können wir machen, indem wir die Identität ausnutzen: $x = e^{\ln(x)}$. Also gilt insbesondere, dass:

$$P(S_n = k) = \frac{1}{\sqrt{2\pi npq}} \left(\frac{k}{np}\right)^{-k} \left(\frac{n-k}{nq}\right)^{-n+k} \Leftrightarrow e^{\ln(P(S_n=k))} \sim e^{\ln\left(\frac{1}{\sqrt{2\pi npq}}\left(\frac{k}{np}\right)^{-k}\left(\frac{n-k}{nq}\right)^{-n+k}\right)}$$

Beschreibe kurz, was genau die obere Gleichung aussagt und wieso wir einfach die e -Funktion und den (natürlichen) Logarithmus anwenden können!

Im Folgenden betrachten wir erst mal nur den Ausdruck im $\ln()$ welchen wir mithilfe der Logarithmengesetze umformen können:

$$\ln \left(\left(\frac{k}{np} \right)^{-k} \left(\frac{n-k}{nq} \right)^{-n+k} \right) = (-np - z\sqrt{npq}) \ln \left(1 + z\sqrt{\frac{q}{np}} \right) + (-nq + z\sqrt{npq}) \ln \left(1 - z\sqrt{\frac{p}{nq}} \right)$$

Nenne die Rechengesetze und Substitutionen, welche vermutlich für diese Umformung genutzt wurden!

Diese Darstellung des Termes sieht zwar erst mal größer aus, hat aber den Vorteil, dass er aus zwei sich sehr ähnlich aussehenden Teilen besteht. Das machen wir uns zu Nutzen! Mithilfe der Taylor-Approximation für den Logarithmus lassen sich nämlich folgende Abschätzungen beweisen:

$$\ln(1 - \alpha) \sim -\alpha - \frac{\alpha^2}{2}$$

$$\ln(1 + \alpha) \sim \alpha - \frac{\alpha^2}{2}$$

Mithilfe einiger Umformungen zufallen dadurch die sich ähnlich sehenden Terme, was zu folgendem, erstaunlichem, Resultat führt:

$$(-np - z\sqrt{npq}) \ln \left(1 + z\sqrt{\frac{q}{np}} \right) + (-nq + z\sqrt{npq}) \ln \left(1 - z\sqrt{\frac{p}{nq}} \right) \sim -\frac{z^2}{2}$$

Zusatzaufgabe:

Rechne dieses Resultat nach!

Setzen wir dieses Ergebnis nun wieder in unsere letzte Version von $P(S_n = k)$ ein, erhalten wir das, was wir beweisen wollten:

$$P(S_n = k) \sim \frac{1}{\sqrt{2\pi npq}} e^{-\frac{z^2}{2}}$$

$$=$$
$$=$$
$$=$$
$$= \frac{1}{\sqrt{2\pi\sigma^2}} \cdot e^{-\frac{(x-\mu)^2}{2\sigma^2}}$$

Daher konvergiert eine Binomialverteilte Zufallsvariable (für $n \rightarrow \infty$) gegen die Dichtefunktion der Normalverteilung.