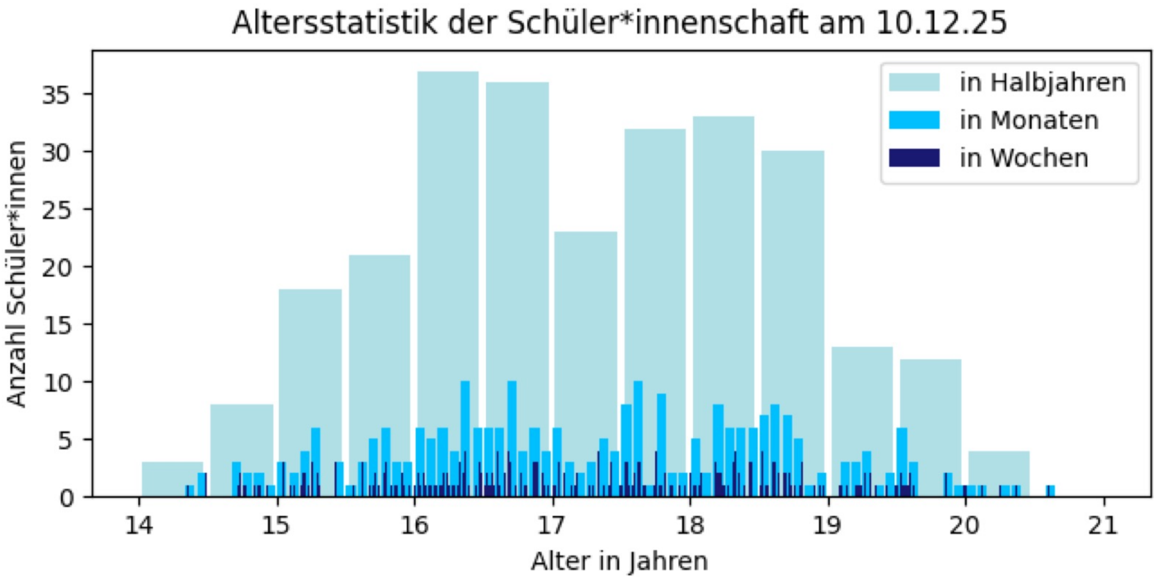
Der Übergang von diskreten zu kontinuierlichen Variablen

Ein paar Vorüberlegungen nach Griffiths

Im QM-Buch von Griffiths stellt der Autor in Abschnitt 1.3.2 Überlegungen an, wie die Statistik der Verteilung einer **diskreten** Variable in eine Statistik der Verteilung einer kontinuierlichen Variable übergeführt werden kann. Dieser Gedankengang sei hier nochmals wiedergegeben und ein wenig ergänzt.

> **Griffiths:** *Bis jetzt sind wir stillschweigend davon ausgegangen, dass wir uns nur mit* **diskreten** *Variablen beschäftigen – also mit Variablen, die nur bestimmte, einzelne Werte annehmen können (in dem Beispiel (der Altersverteilung) musste j eine ganze Zahl sein, da ich das Alter nur in Jahren angegeben hatte). Man kann das aber ganz leicht auf* **kontinuierliche** *Variable verallgemeinern. Wenn Sie eine beliebige Person zufällig auf der Straße auswählen und die Wahrscheinlichkeit betrachten, dass sie* **genau** *16 Jahre, 4 Stunden, 27 Minuten und 3.3333... Sekunden alt ist, dann finden Sie den Wert* **null**.

Diese Aussage möchte ich noch etwas anders veranschaulichen, als Griffiths dies im Folgenden macht. Die folgende Grafik zeigt die Altersverteilung der Schüler∗innen am Gymnasium Unterstrass für den Stichtag 5.11.2023. Die Verteilung wird dreimal gezeigt, nämlich für Halbjahres-, Monats- und Wochenunterteilungen. Klar: Je feiner die Unterteilung ist, desto weniger Schüler∗innen enthält die einzelne Unterteilungsbox. Somit fällt die Verteilung zusammen und die Verteilungswerte nähern sich **null** an. Dies ist also eine zweite, klärende Sichtweise auf die Sache.



> **Griffiths:** *Die einzig vernünftige Aussage betrifft die Wahrscheinlichkeit, dass das Alter der Person in einem bestimmten* **Intervall** *liegt – beispielsweise zwischen 16 und 17. Wenn das Intervall hinreichend kurz ist, dann ist diese Wahrscheinlichkeit* **proportional zur Intervalllänge**. *Beispielsweise ist die Wahrscheinlichkeit, dass die Person zwischen genau 16 Jahren und 16 plus* **zwei** *Tagen alt ist, vermutlich doppelt so hoch wie die Wahrscheinlichkeit für ein Alter zwischen 16 Jahren und 16 plus* **einem** *Tag. (Zumindest dann, wenn es nicht einen außerordentlichen Babyboom vor 16 Jahren, und zwar genau an dem Tag, gegeben hat; in diesem Fall haben wir einfach das Intervall zu lang gewählt, als dass wir die Regel anwenden könnten. Wenn der Babyboom genau sechs Stunden gedauert hat, dann müssen wir eben Intervalle von einer Sekunde Länge oder noch kürzer wählen, um auf der sicheren Seite zu sein. Technisch sprechen wir von* **infinitesimalen** *Intervallen.) Somit gilt:*
>
> $$\left\{ \begin{array}{l} \text{Wahrscheinlichkeit dafür, dass eine zufällig} \\ \text{gewählte Person zwischen } x \text{ und } (x + \mathrm{d}x) \text{ liegt} \end{array} \right\} = \varrho(x)\,\mathrm{d}x$$
>
> *Man könnte den Proportionalitätsfaktor $\varrho(x)$ salopp " die Wahrscheinlichkeit, x zu erwischen" nennen; fachlich korrekt ist die Bezeichnung* **Wahrscheinlichkeitsdichte**. *Die Wahrscheinlichkeit, dass x zwischen a und b liegt, also in einem* **endlichen** *(beschränkten) Intervall, wird mit dem Integral über $\varrho(x)$ angegeben:*
>
> $$P_{ab} = \int_a^b \varrho(x)\,\mathrm{d}x$$

Die Transformation der Regeln von diskreten zu kontinuierlichen Variablen

Griffiths: *Die Regeln, die wir für diskrete Verteilungen hergeleitet haben, gelten in ganz ähnlicher Weise auch hier für kontinuierliche Variable.*

- $P(j)$ = Wahrsch. für den Fall $j \quad \longrightarrow \quad \varrho(x)$ = Wahrsch.dichte für die Variable x

$$P(j) = \frac{N(j)}{N} \qquad \longrightarrow \qquad \varrho(x)\,\mathrm{d}x = \left\{ \begin{array}{l} \text{Wahrscheinlichkeit für einen} \\ \text{Variablenwert zwischen } x \text{ und } x + \mathrm{d}x \end{array} \right\}$$

- **Normierung:** Die Gesamtwahrscheinlichkeit aller Fälle/Variablenwerte beträgt $100\,\% = 1$:

$$\sum_j P(j) = 1 \qquad \longrightarrow \qquad \int_{-\infty}^{\infty} \varrho(x)\,\mathrm{d}x = 1$$

- **Erwartungswert der Verteilung** = Durchschnitt resp. Mittelwert aller Fälle j resp. Werte x:

$$\langle j \rangle = \sum_j j P(j) \qquad \longrightarrow \qquad \langle x \rangle = \int_{-\infty}^{\infty} x \varrho(x)\,\mathrm{d}x$$

- Erwartungswert der Quadrate:

$$\langle j^2 \rangle = \sum_j j^2 P(j) \qquad \longrightarrow \qquad \langle x^2 \rangle = \int_{-\infty}^{\infty} x^2 \varrho(x)\,\mathrm{d}x$$

- Die mittlere Abweichung vom Mittelwert verschwindet stets:

$$\langle \Delta j \rangle = \sum_j \big(j - \langle j \rangle\big) P(j) = 0 \qquad \longrightarrow \qquad \langle \Delta x \rangle = \int_{-\infty}^{\infty} \big(x - \langle x \rangle\big) \varrho(x)\,\mathrm{d}x = 0$$

- **Varianz** σ^2 = mittlere quadratische Abweichung vom Mittelwert ist ein Maß für die Streuung:

$$\sigma^2 = \langle (\Delta j)^2 \rangle = \sum_j (j - \langle j \rangle)^2 P(j) \qquad \longrightarrow \qquad \sigma^2 = \langle (\Delta x)^2 \rangle = \int_{-\infty}^{+\infty} \big(x - \langle x \rangle\big)^2 \varrho(x)\,\mathrm{d}x$$

  Für die Varianz gilt sowohl bei diskreten, wie auch bei kontinuierlichen Variablen:

$$\sigma^2 = \langle j^2 \rangle - \langle j \rangle^2 \qquad \longrightarrow \qquad \sigma^2 = \langle x^2 \rangle - \langle x \rangle^2$$

- **Standardabweichung** σ = Quadratwurzel aus der Varianz. Ebenfalls ein Streuungsmaß:

$$\sigma = \sqrt{\sigma^2} = \sqrt{\langle j^2 \rangle - \langle j \rangle^2} \qquad \longrightarrow \qquad \sigma = \sqrt{\sigma^2} = \sqrt{\langle x^2 \rangle - \langle x \rangle^2}$$

  Typischerweise sind im Bereich (Mittelwert $\pm\sigma$) etwa $68\,\%$ aller Werte enthalten, in (Mittelwert $\pm 2\sigma$) sind es bereits $95\,\%$ und in (Mittelwert $\pm 3\sigma$) sogar über $99\,\%$.

- **Erwartungswert irgendeiner Funktion f der Variable x:**

$$\langle f(j) \rangle = \sum_j f(j) P(j) \qquad \longrightarrow \qquad \langle f(x) \rangle = \int_{-\infty}^{\infty} f(x) \varrho(x)\,\mathrm{d}x$$

  Zu jeder solchen Funktion kann auch eine Varianz resp. eine Standardabweichung berechnet werden:

$$\sigma_f^2 = \langle f^2(j) \rangle - \langle f(j) \rangle^2 \qquad \longrightarrow \qquad \sigma_f^2 = \langle f^2(x) \rangle - \langle f(x) \rangle^2$$