Ergänzungsblatt zur Programmierung von Aufgabe 2.2 b) - e):

- Allgemeiner Tipp für C++: Wenn Variablen erzeugt werden, enthalten sie erst einmal irgendwelche Werte. Wenn diese Variablen zur Summation verwendet werden, **müssen** sie vorher initialisiert werden!
- 2-D Vektoren \underline{x} und \underline{y} können jeweils als Matrix (1, 2) bzw. Matrix (2, 1) erzeugt werden, oder als (x_1, x_2) bzw. (y_1, y_2) Paar. Andere Varianten sind natürlich auch möglich.
- Die Anzahl S von Werten der Zufallsvariablen Y sollte im Bereich von $100-100\,000$ liegen.
- Das Histogramm soll als 2D-Array (z.B. als Matrix) mit den Dimensionen von x_1 und x_2 erzeugt werden. Dabei ist darauf zu achten, dass die Indizes für die Matrix-Elemente nicht mit den Grenzen der Histogramm-Bins verwechselt werden. Das Histogramm **muss nicht** gezeichnet werden.
- Zur Schätzung von Mittelwert und Standardabweichungen sollten die in der Einführungsveranstaltung vorgestellten Formeln implementiert werden. Die geschätzen Mittelwerte $\hat{\mu}$ sollten in der Nähe der in Aufgabe 2.1a) berechneten $\mathcal{E}(X_1)=0$ bzw. $\mathcal{E}(X_2)=-1.0$ liegen, die Varianzen $\hat{\sigma}_1^2\approx 0.033$ und $\hat{\sigma}_2^2\approx 0.133$ entsprechen etwa $\frac{1}{N}$ mal den **Varianzen** aus Aufgabe 2.2a).
- Zur Berechnung des quadratischen Fehlers sollten die Zentren der Histogramm-Bins als <u>z</u>-Werte für die geschätzte Normalverteilung benutzt werden. Als Fehlerfunktion kann der Summed Squared Error

$$E^{SSE} = \sum_{\underline{z}_{bin}} \left[\mathcal{N}(\underline{z}_{bin}) - H(\underline{z}_{bin}) \right]^2$$

verwendet werden, wobei H das Histogramm aus d) bezeichnet. Eine Normierung des Fehlers mit der Anzahl der Bins des Histogramms kann sinnvoll sein, d.h. die Berechnung des *Mean Summed Squared Error* E^{MSSE} .

• Bei genügend großem S sollten die Fehlerwerte in der Größenordnung von $E^{SSE}\approx 0.2\pm 0.1$ bzw. $E^{MSSE}\approx 0.0004\pm 0.0002$ liegen.