a) **► Gleichung der Regressionsgeraden bestimmen** (10P)

Der Operator „bestimmen" lässt einen Einsatz des GTR zu. Du kannst die Gleichung der Regressionsgeraden also mit dem GTR berechnen. Gib hierzu zunächst die Daten in den GTR ein und führe anschließend die lineare Regression durch. Achte darauf, dass die Regressionsgerade die Massen in Abhängigkeit von den Längen der Zylinder angeben soll. Wähle also für die Liste der x-Werte die Zeile zur Länge und als Liste für die y-Werte die Zeile zur Masse.

► Regressionsgerade auswählen und begründen

Es stehen die Geraden g_1 und g_2 als mögliche Regressionsgeraden zur Verfügung, wobei als Grundlage nur die Daten der Zylinder 3, 4 und 5 verwendet wurden. Überlege dir, was die Eigenschaft und die Funktion einer Regressionsgeraden ist:

- Mit einer Regressionsgeraden wird versucht, die **Datenpaare** so gut wie möglich zu **anzunähern**. Es wird also die Gerade gesucht, die von den Punkten, welche die Datenpaare darstellen, **immer möglichst wenig entfernt** ist.
- Die Regressionsgerade unter g_1 und g_2 ist also diejenige, welche die Datenpaare besser annähert.

Du kannst also so vorgehen:

- Berechne die y-Werte der Geraden g_1 und g_2 für die Zylinder 3, 4 und 5.
- Vergleiche die Werte der Geraden mit den echten Werten. „Vergleichen" heißt dabei: Berechne die **Summe der Quadrate der Differenzen**.
- Je kleiner die Summe der Quadrate der Differenzen ist, desto besser die Regressionsgerade.

b) **► Vertrauensintervall bestimmen** (11P)

Sei X zunächst die Anzahl der Zylinder der Qualitätsstufe I in der Stichprobe. X kann näherungsweise als **binomialverteilte** Zufallsgröße angenommen werden mit $n = 200$ und p unbekannt. Einen ersten Schätzwert für p kannst du über die Angabe ermitteln, dass 80 von 200 Zylindern der Qualitätsstufe I zugeordnet wurden, d.h.:
$\frac{80}{200} = 0,4.$
Gesucht ist nun ein Intervall, in dem der tatsächliche Anteil p der Wähler mit einer Wahrscheinlichkeit von 95 % liegt. Einen Ansatz für dieses Problem bieten die σ-Regeln. Diese dürfen angewandt werden, wenn das Laplace-Kriterium $\sigma > 3$ erfüllt ist. Tatsächlich ergibt sich z.B. mit dem Schätzwert 0,4 für p die Standardabweichung
$\sigma = \sqrt{200 \cdot 0,4 \cdot (1-0,4)} \approx 6,93 > 3.$
Selbstverständlich kann dies nur als Näherung gesehen werden. Tendenziell kann aber davon ausgegangen werden, dass die Bedingung $\sigma > 3$ erfüllt ist.

Du kannst also so vorgehen:

- Wähle die σ-Regel, welche eine Aussage über ein 95 %-Konfidenzintervall um den **Erwartungswert** μ macht.
- Bedenke: $\mu = n \cdot p$. Forme den Ausdruck in der σ-Regel also so um, dass er eine Aussage über p macht. Hieraus ergibt sich: $P\left(\left|\frac{X}{n} - p\right| \leq 1{,}96\sqrt{\frac{p \cdot (1-p)}{n}}\right) \leq 0{,}95$.
- Löse die Ungleichung nach p auf und berechne so die Grenzen des Intervalls.

▶ Aussagen begründet entscheiden

Betrachte deinen Ansatz, den du zur Berechnung der Intervallgrenzen verwendet hast:

$$|0{,}4 - p| \leq 1{,}96\sqrt{\frac{p \cdot (1-p)}{200}}$$

Überlege, an welcher Stelle der Stichprobenumfang n und an welcher Stelle die Sicherheitswahrscheinlichkeit auftreten und betrachte, welchen Einfluss sie auf die Länge des Intervalls haben. Dabei ist wichtig: der Ausdruck $1{,}96\sqrt{\frac{p\cdot(1-p)}{200}}$ ist der **Radius** des Intervalls. Je größer der Radius, desto größer ist auch die Länge.

c) **▶ Wahrscheinlichkeit für richtige Entscheidung berechnen** (9P)

Du kannst die Situation in einem Baumdiagramm grafisch darstellen und dir so einen besseren Überblick über die Zusammenhänge verschaffen. Wir wollen dabei auf der ersten Stufe die Unterscheidung „Qualitätsklasse I – nicht Qualitätsklasse I“ machen und auf der zweiten Stufe die Unterscheidung „Qualitätssiegel – kein Qualitätssiegel“.

Auf der ersten Stufe ist dabei die Wahrscheinlichkeit für „Qualitätsklasse I“ genau 0,4. Also ist ein Zylinder mit einer Wahrscheinlichkeit von 0,6 nicht Qualitätsklasse I.

Auf der zweiten Stufe hat ein Zylinder der Qualitätsklasse I mit einer Wahrscheinlichkeit von 90 % ein Qualitätssiegel . Also hat ein Zylinder der Qualitätsstufe I mit einer Wahrscheinlichkeit von 10 % das Qualitätssiegel nicht.

Zugleich hat ein Zylinder, der nicht Qualitätsklasse I ist, mit Wahrscheinlichkeit 0,05 ein Qualitätssiegel; d.h. mit Wahrscheinlichkeit 0,95 hat solch ein Zylinder kein Qualitätssiegel.

Berechne nun die Wahrscheinlichkeit dafür, dass das Prüfgerät eine richtige Entscheidung trifft. Dies ist dann der Fall, wenn es einem Zylinder der Qualitätsklasse I ein Qualitätssiegel gibt **oder** wenn es einem Zylinder von geringerer Qualität kein Qualitätssiegel gibt.

▶ Wahrscheinlichkeit für Zylinder der Klasse I berechnen

Nun wird der Fall betrachtet, dass ein Zylinder mit einem Qualitätssiegel versehen wurde. Du sollst die Wahrscheinlichkeit dafür berechnen, dass der Zylinder **dann** zur Qualitätsklasse I gehört. Gesucht ist also die Wahrscheinlichkeit dafür, dass der Zylinder zur Qualitätsklasse I gehört, **unter der Bedingung dass** er mit einem Siegel versehen wurde.

Es handelt sich hierbei also um eine bedingte Wahrscheinlichkeit. Für diese gilt:

$$P(A \mid B) = \frac{P(A \cap B)}{P(B)}.$$

In unserem Fall ist das Ereignis A: „Ein Zylinder ist Qualitätsklasse I“ und das Ereignis B ist „Ein Zylinder hat ein Siegel“.

Das zweite Ereignis setzt sich dabei aus zwei Ereignissen zusammen:

- Ein Zylinder ist Qualitätsklasse I und hat ein Siegel.
- Ein Zylinder ist nicht Qualitätsklasse I und hat ein Siegel.

Die benötigten Wahrscheinlichkeiten findest du im Baumdiagramm.