a) ▶ **Wahrscheinlichkeiten berechnen** (11P)

X beschreibe die Anzahl der Zigarettenschachteln ohne Steuerbanderole; X kann als binomialverteilt angenommen werden:

- Eine Schachtel hat eine Steuerbanderole oder nicht
- Die Wahrscheinlichkeit $p = 0{,}107$ wird unabhängig von Schachtel zu Schachtel als gleichbleibend angenommen

Die Parameter für die Binomialverteilung sind $n = 40$ und $p = 0{,}107$.

Gefragt ist zunächst nach der Wahrscheinlichkeit $P(X = 4)$.

$$P(X=4) = \binom{40}{4} \cdot (0{,}107)^4 \cdot (0{,}893)^{36} = 0{,}2037$$

Mit einer Wahrscheinlichkeit von 20,37 % besitzen 4 der 40 Zigarettenschachteln keine Steuerbanderole. Nun ist zu berechnen: $P(3 \leq X \leq 5)$.

▶▶ **Lösungsweg A: Lösung von Hand**

Aufsummieren der Einzelwahrscheinlichkeiten $P(X = k)$ ergibt:

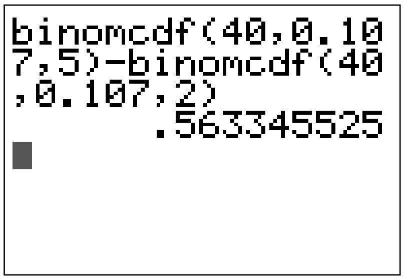
$$P(3 \leq X \leq 5) = \sum_{k=3}^{5} P(X=k) = \sum_{k=3}^{5} \binom{40}{k} \cdot (0{,}107)^k \cdot (0{,}893)^{40-k} = 0{,}5633$$

Mit einer Wahrscheinlichkeit von 56,33 % haben mindestens und höchstens 5 der 10 Zigarettenschachteln keine Steuerbanderole.

▶▶ **Lösungsweg B: Lösung mit dem GTR**

Es ist $P(3 \leq X \leq 5) = P(X \leq 5) - P(X \leq 2)$.

Berechne die Wahrscheinlichkeit mit `2nd → VARS (DISTR) → binomcdf(`. Achte auf die Reihenfolge: `binomcdf(n,p,k)`.



Mit einer Wahrscheinlichkeit von 56,33 % haben mindestens und höchstens 5 der 10 Zigarettenschachteln keine Steuerbanderole.

▶ **Mindestanzahl berechnen**

Sei X wieder die Anzahl der Zigarettenschachteln ohne Steuerbanderole. X ist binomialverteilt mit n unbekannt und $p = 0{,}107$.

n soll bestimmt werden unter der Bedingung, dass $P(X \geq 1) \geq 0{,}8$:

$$\begin{aligned}
P(X \geq 1) &\geq 0{,}8 \\
1 - P(X = 0) &\geq 0{,}8 \\
1 - \binom{n}{0} \cdot (0{,}107)^0 \cdot (0{,}893)^n &\geq 0{,}8 \\
1 - (0{,}893)^n &\geq 0{,}8 \\
-(0{,}893)^n &\geq -0{,}2 \\
(0{,}893)^n &\leq 0{,}2 && \mid \ln(\,) \\
n \cdot \ln(0{,}893) &\leq \ln(0{,}2) && \mid : \ln(0{,}893) \quad \text{Achtung: } \ln(0{,}893) < 0 \\
n &\geq \frac{\ln(0{,}2)}{\ln(0{,}893)} \approx 14{,}22
\end{aligned}$$

Es müssen mindestens 15 Schachteln eingesammelt werden, damit mit einer Wahrscheinlichkeit von mindestens 80 % mindestens eine keine Steuerbanderole hat.

b) ▶ **Wahrscheinlichkeit für Beschlagnahmung berechnen** (10P)

Die Lieferung besteht aus insgesamt 100 Stangen, von denen 8 unverzollt sind. Der Anteil unverzollter Stangen liegt also bei $p = \frac{8}{100}$. Die Polizei entnimmt nun **nacheinander** fünf Stangen; es handelt sich um ein **Ziehen ohne Zurücklegen**.

Die Lieferung wird beschlagnahmt, wenn **eine** unverzollte Stange gefunden wird. Nach der Pfadregel ergibt sich dafür die Wahrscheinlichkeit:

$$\begin{aligned} P\,(\text{mindestens eine unverzollte Stange}) &= 1 - P\,(\text{keine unverzollte Stange}) \\ &= 1 - \frac{92}{100} \cdot \frac{91}{99} \cdot \frac{90}{98} \cdot \frac{89}{97} \cdot \frac{88}{96} \\ &= 1 - \frac{26611}{40740} \approx 1 - 0{,}6532 = 0{,}3468 \end{aligned}$$

Mit einer Wahrscheinlichkeit von 34,68 % wird die Lieferung beschlagnahmt.

▶ **Höchstanzahl unverzollter Stangen ermitteln**

Sei k zunächst die unbekannte Anzahl unverzollter Stangen in der Lieferung. Die Polizei entnimmt wieder nacheinander 5 Stangen. Nach der Überlegung aus (1) folgt für die Wahrscheinlichkeit $P(\text{keine unverzollte Stange})$:

$$P(\text{keine unverzollte Stange}) = \frac{100-k}{100} \cdot \frac{99-k}{99} \cdot \frac{98-k}{98} \cdot \frac{97-k}{97} \cdot \frac{96-k}{96}.$$

Diese Wahrscheinlichkeit soll größer als 90 % sein. Systematisches Probieren liefert:

$$k = 3: \quad \frac{97}{100} \cdot \frac{96}{99} \cdot \frac{95}{98} \cdot \frac{94}{97} \cdot \frac{93}{96} \approx 0{,}856$$

$$k = 2: \quad \frac{98}{100} \cdot \frac{97}{99} \cdot \frac{96}{98} \cdot \frac{95}{97} \cdot \frac{94}{96} \approx 0{,}902$$

Es folgt: Es dürfen höchstens zwei unverzollte Stangen in der Lieferung sein, damit mit einer Wahrscheinlichkeit von mehr als 90 % keine in der Stichprobe gefunden wird.

c) ▶ **Anzahl unverzollter Schachteln schätzen** (8P)

In den öffentlichen Mülltonnen der Stadt wurden insgesamt 200 Schachteln entsorgt. In der Stichprobe, die in der Entsorgungsstation durchgeführt wurde, wurden von diesen 200 Schachteln genau 14 Stück gefunden, d.h. $\frac{14}{200} = \frac{7}{100} = 7\,\%$.

In einem einfachen Modell können wir davon ausgehen, dass die Abfälle der Stadt „gut gemischt“ und die Stichprobe deshalb aussagekräftig ist. Dies lässt den Schluss zu, dass auch die 146 unverzollten Schachteln jeweils **7 %** der gesamten Menge darstellen.

Für die Anzahl insgesamt angelieferter unverzollter Schachteln ergibt sich dann nach dem einfachen Dreisatz:

$$7\,\% \;\widehat{=}\; 146 \quad\Leftrightarrow\quad 1\,\% \;\widehat{=}\; 20{,}85 \quad\Leftrightarrow\quad 100\,\% \;\widehat{=}\; 2085$$

An diesem Tag wurden schätzungsweise 2085 unverzollte Zigarettenschachteln in der Entsorgungsstation angeliefert.

d) ▸ **Konfidenzintervall bestimmen** (21P)

Wir sollen ein 90 %-Konfidenzintervall für p bestimmen, wobei p der noch unbekannte Anteil unverzollter Schachteln ist.

Hierzu wurde eine Stichprobe von 200 Schachteln ausgewertet. Sei X die Anzahl der unverzollten Schachteln in der Stichprobe. X kann als binomialverteilt angenommen werden mit p unbekannt und $n = 200$.

Als Ansatz für das Konfidenzintervall bedienen wir uns der σ-Regeln:

$P\left(\mu - 1{,}64\sigma < X < \mu + 1{,}64\sigma\right) \approx 0{,}9$

bzw. mit $\frac{X}{n} = \dfrac{22}{100} = 0{,}11$:

$P\left(\frac{\mu}{n} - 1{,}64\frac{\sigma}{n} < 0{,}11 < \frac{\mu}{n} + 1{,}64\frac{\sigma}{n}\right) \approx 0{,}9.$

Dabei gilt $\mu = n \cdot p = 200p$ und $\sigma = \sqrt{n \cdot p \cdot (1-p)} = \sqrt{200p - 200p^2}$

Einsetzen in die Ungleichung aus der σ-Regel ergibt:

$$\begin{aligned}
\frac{200p}{200} - 1{,}64\frac{\sqrt{200p-200p^2}}{200} &< 0{,}11 < \frac{200p}{200} + 1{,}64\frac{\sqrt{200p-200p^2}}{200} \\
p - 1{,}64\frac{\sqrt{200p-200p^2}}{200} &< 0{,}11 < p + 1{,}64\frac{\sqrt{200p-200p^2}}{200} \qquad \mid -p \\
-1{,}64\frac{\sqrt{200p-200p^2}}{200} &< 0{,}11 - p < 1{,}64\frac{\sqrt{200p-200p^2}}{200}
\end{aligned}$$

Die letzte Ungleichung ist äquivalent zu $|0{,}11 - p| < 1{,}64\dfrac{\sqrt{200p-200p^2}}{200}$.

Wir suchen das **kleinstmögliche** p und können deshalb mit „=“ weiter rechnen:

$$\begin{aligned}
|0{,}11 - p| &= 1{,}64\frac{\sqrt{200p-200p^2}}{200} \qquad \mid (\,)^2 \\
(0{,}11 - p)^2 &= \frac{1{,}64^2}{200^2} \cdot (200p - 200p^2) \\
0{,}11^2 - 0{,}22p + p^2 &= \frac{1{,}64^2}{200} \cdot p - \frac{1{,}64^2}{200} \cdot p^2 \\
p^2 \cdot \left(1 + \frac{1{,}64^2}{200}\right) - p \cdot \left(0{,}22 + \frac{1{,}64^2}{200}\right) + 0{,}11^2 &= 0 \\
1{,}01345 \cdot p^2 - 0{,}233448 \cdot p + 0{,}0121 &= 0 \\
p^2 - 0{,}23035 \cdot p + 0{,}011939 &= 0
\end{aligned}$$

$$\begin{aligned}
p_{1,2} &= 0{,}115175 \pm \sqrt{0{,}115175^2 - 0{,}011939} \\
p_{1,2} &= 0{,}115175 \pm \sqrt{0{,}001326} \\
p_{1,2} &= 0{,}115175 \pm 0{,}036414
\end{aligned}$$

Daraus folgen die Werte $p_1 = 0{,}151589$ und $p_2 = 0{,}078761$ und schlussendlich das Konfidenzintervall $[0{,}079; 0{,}1516]$.

▸ **Ungleichung beweisen**

$p \cdot (1-p) \leq 0{,}2 \cdot 0{,}8 \;\Leftrightarrow\; p - p^2 \leq 0{,}16 \;\Leftrightarrow\; p^2 - p + 0{,}16 \geq 0$

Betrachten wir die Ungleichung: Du kannst dir $p^2 - p + 0{,}16$ als Funktionsterm einer nach oben geöffneten Parabel vorstellen.
Gesucht werden die Intervalle, in denen sie oberhalb der x-Achse verläuft. Da sie nach oben geöffnet ist, haben diese Intervalle die Form $I_1 =]-\infty; p_1[$ und $I_2 =]p_2; \infty[$. Damit erhalten wir als Lösung nach der p-q-Formel:

$$p_1 \geq \frac{1}{2} + \sqrt{\frac{1}{4} - 0{,}16} \Leftrightarrow p_1 \geq 0{,}8 \qquad \text{und} \qquad p_1 \geq \frac{1}{2} - \sqrt{\frac{1}{4} - 0{,}16} \Leftrightarrow p_2 \leq 0{,}2.$$

Damit ist die Ungleichung nachgewiesen.

▶ Stichprobenumfang näherungsweise bestimmen

Eine Schätzung gilt als hinreichend genau, wenn sie mit einer Wahrscheinlichkeit von mindestens 95 % um höchstens 0,02 vom Anteil p abweicht. Sei X die Anzahl der unverzollten Zigarettenschachteln. X kann als binomialverteilt angenommen werden mit n unbekannt und $p \leq 0{,}2$.
Aus den σ-Regeln folgt wieder Ansatz:

$$P(\mu - 1{,}96\sigma < X < \mu + 1{,}96\sigma) \approx 0{,}95 \quad \text{bzw.} \quad P\left(\tfrac{\mu}{n} - 1{,}96\tfrac{\sigma}{n} < \tfrac{X}{n} < \tfrac{\mu}{n} + 1{,}96\tfrac{\sigma}{n}\right) \approx 0{,}95$$

Nach der Überlegung aus Aufgabenteil (1) und mit $\frac{\mu}{n} = \frac{n \cdot p}{n} = p$ folgt daraus die Ungleichung $\left|\frac{X}{n} - p\right| \leq 1{,}96 \cdot \frac{\sigma}{n}$.

Dabei stellt $\left|\frac{X}{n} - p\right|$ gerade die **Abweichung** des relativen Anteils $\frac{X}{n}$ in der Stichprobe vom wahren Anteil p. Diese Abweichung soll bei einer hinreichend genauen Schätzung **maximal** 0,02 sein:

$$\begin{aligned}
\left|\frac{X}{n} - p\right| \leq 1{,}96 \cdot \frac{\sigma}{n} &\leq 0{,}02 \\
1{,}96 \cdot \frac{\sigma}{n} &\leq 0{,}02 \\
1{,}96 \cdot \frac{\sqrt{n \cdot p \cdot (1-p)}}{n} &\leq 0{,}02 \\
1{,}96 \cdot \frac{\sqrt{p \cdot (1-p)}}{\sqrt{n}} &\leq 0{,}02
\end{aligned}$$

Da $p \leq 0{,}2$ vorausgesetzt ist, gilt die Ungleichung von oben: $p \cdot (1-p) \leq 0{,}2 \cdot 0{,}8$:

$$\begin{aligned}
1{,}96 \cdot \frac{\sqrt{p \cdot (1-p)}}{\sqrt{n}} \leq 1{,}96 \cdot \frac{\sqrt{0{,}2 \cdot 0{,}8}}{\sqrt{n}} &\leq 0{,}02 \\
1{,}96 \cdot \frac{\sqrt{0{,}2 \cdot 0{,}8}}{\sqrt{n}} &\leq 0{,}02 \\
1{,}96 \cdot \frac{\sqrt{0{,}2 \cdot 0{,}8}}{0{,}02} &\leq \sqrt{n} \\
39{,}2 &\leq \sqrt{n} \\
n &\geq 1536{,}64
\end{aligned}$$

Es müssen mindestens 1537 Zigarettenschachteln untersucht werden, um bei einem tatsächlichen Anteil $p \leq 0{,}2$ eine hinreichend genaue Schätzung möglich zu machen.