a.) (1) ► **Wahrscheinlichkeit, dass genau 2 Personen P - Kranke sind** (6BE)

Betrachtet wird hier die Zufallsvariable X. X beschreibt die Anzahl der P - Kranken Personen in der Bevölkerung und ist eine binomialverteilte Zufallsvariable. Die Wahrscheinlichkeit, dass eine der untersuchten 50 Personen erkrankt ist, entspricht dem relativen Anteil der P - Kranken in der Bevölkerung des untersuchten Gebietes, also also $p = 0,02$.
$\Longrightarrow$ X ist binomialverteilt mit $n = 50$ und $p = 0,02$.

Die Wahrscheinlichkeiten einer Binomialverteilung lassen sich für verschiedene k, wobei k die Anzahl der erkrankten Personen beschreibt, über diesen Term berechnen:
$P(X = k) = \binom{n}{k} \cdot p^k \cdot (1-p)^{n-k}$

Die Wahrscheinlichkeit, dass genau 2 der 50 Personen erkrankt sind, berechnest du nun so über den angegebenen Term:
$P(X = 2) = \binom{50}{2} \cdot 0,02^2 \cdot (1-0,02)^{50-2} = 1225 \cdot 0,0004 \cdot 0,98^{48} \approx 0,186$

Die Wahrscheinlichkeit, dass genau 2 der 50 Personen erkrankt sind, liegt bei $\approx$ 0,186 (18,6 %).

(2) ► **Wahrscheinlichkeit, dass mindestens eine Person P - erkrankt ist**

Betrachtet wird hier die gleiche Zufallsvariable X wie im Aufgabenteil zuvor. Die Wahrscheinlichkeit $P(X \geq 1)$, dass mindestens eine der 50 untersuchten Personen P - erkrankt ist, berechnest du über dessen zugehöriges Gegenereignis. Das Gegenereignis zu mindestens einer erkrankten Person ist:
$\Longrightarrow$ Keine der untersuchten Personen ist P - erkrankt, also $P(X < 1) \;\Leftrightarrow\; P(X = 0)$.

Die gesuchte Wahrscheinlichkeit $P(X \geq 1)$ berechnest du nun wie folgt:

$$\begin{aligned}
P(X \geq 1) &= 1 - P(X < 1) \; = \; 1 - P(X = 0) \\
P(X \geq 1) &= 1 - \left(\binom{50}{0} \cdot 0,02^0 \cdot (1-0,02)^{50-0}\right) \\
P(X \geq 1) &= 1 - (1 \cdot 1 \cdot 0,98^{50}) \\
P(X \geq 1) &= 0,636
\end{aligned}$$

Die Wahrscheinlichkeit, dass mindestens eine der 50 untersuchten Personen P - erkrankt ist, beträgt 0,636 (63,6 %).

b.1) ► **Wahrscheinlichkeit, dass ein Mann P - erkrankt ist**

Gesucht ist die Wahrscheinlichkeit, dass ein Mann P-erkrankt ist. Dies ist gleich der Wahrscheinlichkeit dafür, dass eine zufällig ausgewählte untersuchte Person P-erkrankt ist **unter der Bedingung**, dass sie ein Mann ist.
Du kannst vorgehen:

- Betrachte die Informationen aus der Aufgabenstellung genauer und schreibe die Wahrscheinlichkeiten genau heraus.
- Fertige eine Vierfeldertafel an.
- Bestimme zuletzt die gesuchte Wahrscheinlichkeit.

1. Schritt: Wahrscheinlichkeiten formulieren

Betrachte die Informationen, die dir in der Aufgabenstellung gegeben sind:

- 2 % aller untersuchten Personen weisen die Krankheit auf, d.h. eine zufällig ausgewählte untersuchte Person ist mit einer Wahrscheinlichkeit von 2 % P-erkrankt.
- 49,8 % aller **untersuchten** Personen waren Männer, d.h. eine zufällig ausgewählte untersuchte Person ist mit einer Wahrscheinlichkeit von 49,8 % ein Mann.
- 50 % aller **gesunden** Personen waren Frauen, d.h. eine zufällig ausgewählte untersuchte Person ist, **unter der Bedingung** dass sie gesund ist, eine Frau. Es handelt sich hier also um eine **bedingte** Wahrscheinlichkeit.

Wir wollen zunächst noch die beiden Ereignisse

$K:$ „Eine zufällig ausgewählte untersuchte Person ist P-erkrankt" und

$M:$ „Eine zufällig ausgewählte untersuchte Person ist ein Mann"

definieren. Dabei sollen $\overline{K}$ und $\overline{M}$ die zugehörigen Gegenereignisse sein. Dann können wir obige Wahrscheinlichkeiten so ausdrücken:

$P(K) = 0,02; \qquad P(M) = 0,498 \qquad$ und $\qquad P_{\overline{K}}(\overline{M}) = 0,5.$

Außerdem ist bekannt:

$P(\overline{K}) = 1 - P(K) = 0,98; \qquad P(\overline{M}) = 1 - P(M) = 0,502.$

Betrachte die bedingte Wahrscheinlichkeit von oben genauer:

$$
\begin{aligned}
P_{\overline{K}}(\overline{M}) &= 0,5 && \left|\; P_{\overline{K}}(\overline{M}) = \frac{P(\overline{K} \cap \overline{M})}{P(\overline{K})}\right. \\
\frac{P(\overline{K} \cap \overline{M})}{P(\overline{K})} &= 0,5 && |\; P(\overline{K}) = 1 - P(K) = 0,98 \\
\frac{P(\overline{K} \cap \overline{M})}{0,98} &= 0,5 && |\; \cdot 0,98 \\
P(\overline{K} \cap \overline{M}) &= 0,49
\end{aligned}
$$

2. Schritt: Vierfeldertafel anfertigen

Fassen wir nun alle Wahrscheinlichkeiten, die bisher bekannt sind, in einer Vierfeldertafel zusammen:

	M	$\overline{M}$	$\sum$
K			0,02
$\overline{K}$		0,49	0,98
$\sum$	0,498	0,502	**1**

Vollständiges Ausfüllen der Vierfeldertafel liefert:

	M	$\overline{M}$	$\sum$
K	**0,008**	**0,012**	0,02
$\overline{K}$	**0,49**	0,49	0,98
$\sum$	0,498	0,502	**1**

3. Schritt: Wahrscheinlichkeit berechnen

Gesucht ist die Wahrscheinlichkeit:

$$P_M(K) = \frac{P(M \cap K)}{P(M)} = \frac{0,008}{0,498} \approx 0,016$$

Mit einer Wahrscheinlichkeit von etwa 1,6 % hat ein Mann die P-Krankheit.

b.2) ▶ **Überprüfen auf stochastische Abhängigkeit**

Deine Aufgabe ist es hier zu überprüfen, ob eine Erkrankung mit P vom Geschlecht abhängig ist. Das heißt du überprüfst auf eine stochastische Abhängigkeit zwischen: „Person ist ein Mann“ und „Person ist krank“.
Besteht eine Abhängigkeit zwischen diesen zwei Ereignissen, so besteht eine Abhängigkeit der P - Erkrankung und dem Geschlecht.
Würde stochastische Unabhängigkeit zwischen den genannten Eigenschaften bestehen, so müsste folgendes Produkt einer wahren Aussage entsprechen:

$P(M \cap K) = P(M) \cdot P(K)$

Der Vierfeldertafel ist zu entnehmen: $P(M \cap K) = 0,008$.

$0,008 \neq 0,498 \cdot 0,02 = 0,00996$

Da das Produkt der Wahrscheinlichkeiten $P(M)$ und $P(K)$ nicht dem Eintrag aus der Vierfeldertafel zu $P(M \cap K)$ entspricht, besteht eine stochastische Abhängigkeit zwischen „Person ist ein Mann“ und „Person ist krank“.
$\Longrightarrow$ Das heißt die Erkrankung mit der Krankheit P ist vom Geschlecht der Menschen abhängig.

c.) (1) ▶ **Beschreiben der möglichen Fehlentscheidungen** (5BE)

Folgende zwei Fehlentscheidungen können hier getroffen werden:

(1) Fehler erster Art
Der Fehler erster Art beschreibt jene Situation, in welcher eine Person gesund ist und jedoch mehr als 75 P - Teilchen in seiner Blutprobe gefunden wurden. Das heißt, die untersuchte Person wird fälschlicherweise für krank gehalten.

(2) Fehler zweiter Art
Der Fehler zweiter Art beschreibt jene Situation, in welcher eine Person krank ist und jedoch höchstens 75 P - Teilchen in seiner Blutprobe gefunden wurden. Das heißt, die untersuchte Person wird fälschlicherweise für gesund gehalten.

(2) ▶ **Berechnen der Wahrscheinlichkeiten für diese Fehlentscheidungen**

(1) Fehler erster Art

Betrachtet wird die Zufallsvariable X, welche die Anzahl der gefundenen P - Teilchen im Blut einer gesunden Person beschreibt. X ist binomialverteilt und tritt bei einer gesunden Person mit einer Wahrscheinlichkeit von $p = 0,05$ auf. Zu berechnen ist die Wahrscheinlichkeit dafür, dass sich mehr als $k = 75$ P - Teilchen $(P(X > 75))$, in einer Blutprobe eines gesunden Menschen im Umfang von $n = 1000$ Teilchen, befinden:

$$P(X > 75) = 1 - P(X \leq 75) = 1 - \sum_{i=0}^{75} \binom{1000}{k} \cdot 0,05^k \cdot (1 - 0,05)^{1000-k}$$

Berechne diesen Term mit Hilfe der Summenfunktion deines Taschenrechners oder entnehme die gesuchte Wahrscheinlichkeit der Tabelle für die summierte Binomialverteilung. Für $n = 1000$, $p = 0,05$ und $k = 75$ gilt:

$\Longrightarrow \ P(X > 75) = 1 - P(X \leq 75) = 1 - 0,9997 = 0,0003$

Die Wahrscheinlichkeit dafür, dass eine Person fälschlicherweise für krank gehalten wird, ist $0,0003$ (0,03 %).

(2) Fehler zweiter Art

Hier beschreibt die betrachtete Zufallsvariable Y, die Anzahl der P - Teilchen im Blut einer erkrankten Person. Y ist ebenfalls binomialverteilt und tritt bei einer erkrankten Person mit einer Wahrscheinlichkeit von $p = 0,10$ auf. Zu Berechnen ist hier die Wahrscheinlichkeit dafür, dass sich weniger als $k = 75$ P - Teilchen $(P(X \leq 75))$ in einer Blutprobe einer erkrankten Person, im Umfang von $n = 1000$ Teilchen, befinden:

$$P(X \leq 75) = \sum_{i=0}^{75} \binom{1000}{k} \cdot 0,10^k \cdot (1 - 0,10)^{1000-k}$$

Berechne diesen Term auch hier mit Hilfe der Summenfunktion deines Taschenrechners oder entnehme die gesuchten Wahrscheinlichkeit der Tabelle der summierten Binomialverteilung. Für $n = 1000$, $p = 0,10$ und $k = 75$ gilt:

$\Longrightarrow \ P(X \leq 75) = 0,0038$

Die Wahrscheinlichkeit dafür, dass eine Person fälschlicherweise für gesund gehalten wird, ist $0,0038$ (0,38 %).

d.1) ▶ **Bestimmen der Entscheidungsregel** (11BE)

Es handelt sich hier um einen rechtsseitigen Signifikanztest, bei welchem bestimmt werden soll, ob der relative Anteil der P - erkrankten Personen in der Bevölkerung gestiegen ist. Vor dem Test wurde davon ausgegangen, dass 2 % der Bevölkerung P - erkrankt ist. Da diese Hypothese nun überprüft werden soll, definiert diese die Nullhypothese H_0 des Signifikanztests:

$H_0 : p_0 \leq 0,02$

Die Entscheidungsregeln des Tests werden über den Annahme- und Ablehnungsbereich der Nullhypothese bestimmt. Betrachtet wird hier als Testgröße die Zufallsvariable X, welche die Anzahl der P - Erkrankten in der Bevölkerung repräsentiert. Die Zufallsvariable X ist mit $n = 1000$ und $p = 0,02$ binomialverteilt.

Zu bestimmen sind:

Annahmebereich: $A = \{0, 1, ..., k - 1\}$

Ablehnungsbereich: $\overline{A} = \{k, k + 1, ..., n\}$

Die gesuchte Größe k, welche die obere Grenze des Annahmebereichs, sowie die untere Grenze des Ablehnungsbereichs definiert, bestimmst du mit Hilfe der gegebenen Irrtumswahrscheinlichkeit. Eine Irrtumswahrscheinlichkeit von 10 % gibt an, dass die Wahrscheinlichkeit, dass ein Ereignis aus dem Ablehnungsbereich eintritt höchstens 10 % sein darf.
k ist also über folgende Ungleichung zu bestimmen:

$P(X \geq k) \leq 0,10 \;\Leftrightarrow\; 1 - P(X < k) \leq 0,10 \;\Leftrightarrow\; P(X \leq k-1) \geq 0,90.$

Da die dir gegebene Tabelle zur kumulierten Binomialverteilung für $n = 1000$ nicht ausreicht um k zu bestimmen, approximierst du die Binomialverteilung durch die Normalverteilung.
Bestimme dazu zuerst Erwartungswert μ und Standardabweichung σ der binomialverteilten Zufallsvariablen X:
$\mu = n \cdot p = 1000 \cdot 0,02 = 20$
$\sigma = \sqrt{n \cdot p \cdot (1-p)} = \sqrt{1000 \cdot 0,02 \cdot (1-0,02)} \approx 4,43$

Da $\sigma > 3$ ist, darf der Satz von DeMoivre - Laplace angewandt und k mit Hilfe der Normalverteilung bestimmt werden.
Gehe dabei schrittweise vor:

1. Schritt:
Bestimme mit Hilfe der Tabelle der Normalverteilung jene Stelle z, für die gilt:
$P(X \leq k-1) = \Phi(z) \geq 0,90 \Longrightarrow z \approx 1,29$

2. Schritt:
Transformiere z wie folgt, um dieses auf die Gegebenheiten der Binomialverteilung anzupassen. Vergiss dabei nicht die Stetigkeitskorrektur $(k+0,5)$.

$$z \geq \frac{(k-1+0,5)-\mu}{\sigma} \;\Leftrightarrow\; z \geq \frac{(k-0,5)-20}{4,43}$$

Da dir bekannt ist, dass die Normalverteilung für $z \approx 1,29$ das erste Mal einen Wert größer 0,9 annimmt, kannst du mit Hilfe dieses Wertes und der obigen Transformation von z das gesuchte k bestimmen. Setzte dafür z in die obige Transformation ein und löse nach k auf:

$$\begin{aligned} z &\geq \frac{(k-0,5)-20}{4,43} \\ 1,29 &\geq \frac{(k-0,5)-20}{4,43} && \mid \cdot 4,43 \\ 5,715 &\geq (k-0,5)-20 && \mid +20 \\ 25,715 &\geq k-0,5 \;\Leftrightarrow\; k = 26,215 \end{aligned}$$

Da du k nun bestimmt hast, folgt für den Annahme- bzw. Ablehnungsbereich:

Annahmebereich: $A = \{0, 1, ..., 26\}$
Ablehnungsbereich: $\overline{A} = \{27, 28, ..., n\}$

$\Longrightarrow$ Werden also mehr als 27 P - Kranke in einer Stichprobe von $n = 1000$ Menschen gefunden, so kann davon ausgegangen, dass der Anteil der P - Kranke in der Bevölkerung gestiegen ist.

d.2 ▶ **Wahrscheinlichkeit für den Fehler 2. Art**

Ist der Anteil der P - Kranken tatsächlich auf 3 % gestiegen, so besteht das Risiko, dass die Krankenkasse die Beiträge trotz allem nicht erhöht. Dieser Fall tritt ein, wenn sich in der Stichprobe der Krankenkasse die Anzahl der P - Kranken innerhalb, des in der vorherigen Aufgabe bestimmten, Annahmebereichs bewegt, obwohl sich der Anteil der P - Kranken auf 3 % erhöht hat.

Betrachtet wird hier die binomialverteilte Zufallsvariable X, welche mit $p = 0,03$ und $n = 1000$ verteilt ist. X beschreibt dabei die Anzahl der P - erkrankten in der Stichprobe. Gesucht ist die Wahrscheinlichkeit für den Fehler 2. Art, also: $P(X \leq 26)$.
Da die dir gegebene Tabelle für die kumulierte Binomialverteilung auch hier nicht ausreicht um die Wahrscheinlichkeit zu bestimmen, approximierst du diese wieder mit der Normalverteilung.

Bestimme dazu wieder Erwartungswert μ und Standardabweichung σ der binomialverteilten Zufallsvariablen X:
$\mu = n \cdot p = 1000 \cdot 0,03 = 30$
$\sigma = \sqrt{n \cdot p \cdot (1-p)} = \sqrt{1000 \cdot 0,03 \cdot (1-0,03)} \approx 5,394$

Da $\sigma > 3$ ist, darf der Satz von Moivre - Laplace angewandt und k mit Hilfe der Normalverteilung bestimmt werden. Transformiere dazu wie zuvor z, um diese auf die Gegebenheiten der Binomialverteilung anzupassen:

$$z = \frac{(k+0,5)-\mu}{\sigma} \Leftrightarrow z = \frac{(k+0,5)-30}{5,394}$$

Mit $k = 26$ folgt:

$$z = \frac{(26+0,5)-30}{5,394} \Leftrightarrow z = -0,649$$

Bestimme nun mit Hilfe der Tabelle zur Normalverteilung die gesuchte Wahrscheinlichkeit $P(X \leq 26)$, in dem du folgenden Satz anwendest:

$$P(X \leq 26) = \Phi(-0,649) = 1 - \Phi(0,649) = 1 - 0,7422 = 0,259$$

$\Longrightarrow$ Die Wahrscheinlichkeit, dass die Krankenkasse die Beiträge fälschlicherweise nicht erhöht liegt bei 0,259 (25,9 %).