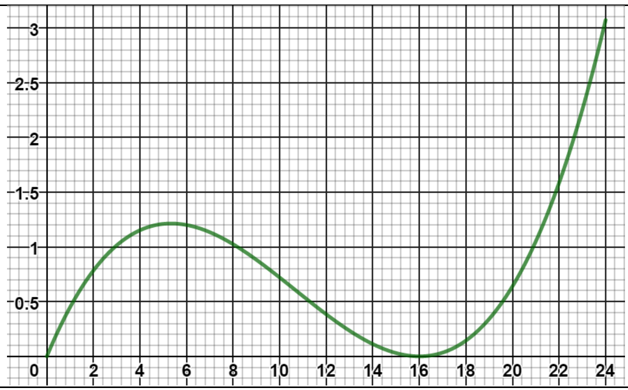
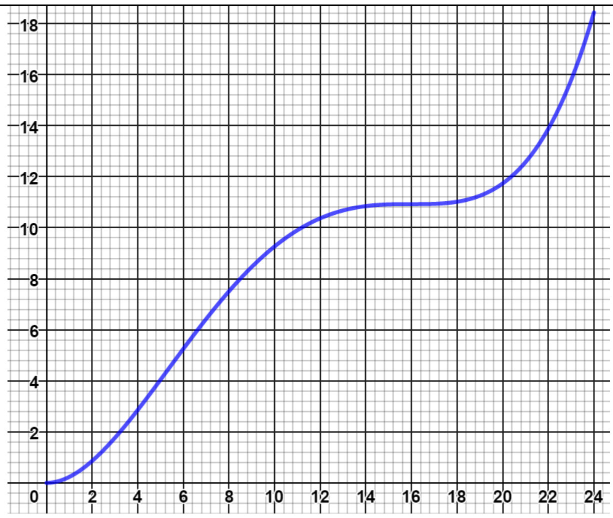
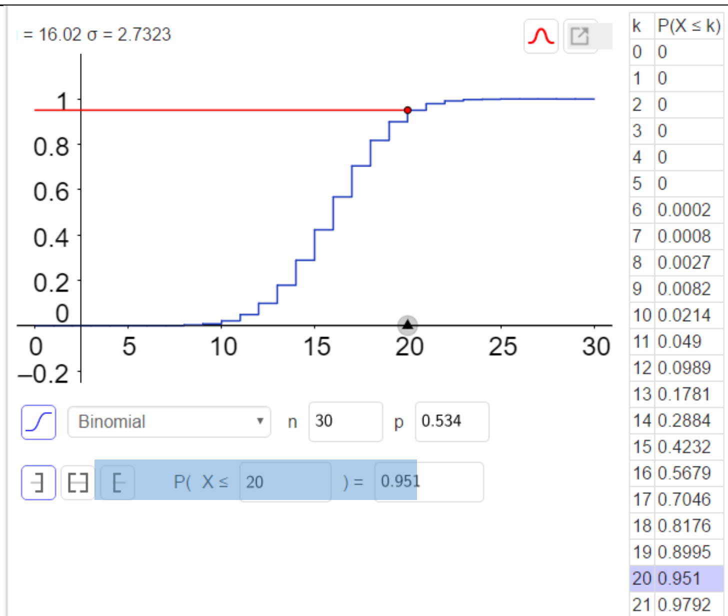
## Teil B Antworten

| | Teil B2 | Punkte KV | Punkte M | Punkte P | Punkte I |
|---|---|---|---|---|---|
| | Im Jahr 2002 wurden im Großherzogtum Luxemburg die monatlichen Durchschnittstemperaturen aufgezeichnet. Es ist bekannt, dass der Januar 2002 mit 1,6°C der kälteste Monat war und die höchste Durchschnittstemperatur im Juni 2002 mit 18,6°C gemessen wurde. | | | | |
| a) | **Begründen** Sie, dass in Europa die monatlichen Durchschnittstemperaturen für einige aufeinanderfolgende Jahre mit einem periodischen Modell modelliert werden können. | | | | |
| | Die Temperaturen steigen und fallen innerhalb eines Jahres und dieser Vorgang wiederholt sich jedes Jahr. Obwohl die Temperaturen nicht jedes Jahr gleich sind, sind im Durchschnitt die Sommermonate warm und die Wintermonate kalt. | 2 | | | |
| b) | **Geben Sie** die Amplitude und die Periode dieses Modells **an**. | | | | |
| | Amplitude $a = 8{,}5$<br>Zeitraum $T = 12$ | 1 | 1 | | |
| c) | **Bestimmen Sie** die Parameter $a$, $b$, $c$ und $d$ in dem Modell $T$ mit<br>$T(x) = a \cdot sin(b \cdot (x - c)) + d$<br>das die gegebenen Daten beschreibt, wobei $x$ der Monat ist, beginnend mit $x = 1$ für Januar 2002 und $T$ die durchschnittliche Temperatur ist. | | | | |
| | $a = \frac{18{,}6 - 1{,}6}{2} = 8{,}5$<br>$d = 8{,}5 + 1{,}6 = 10{,}1$<br>$b = \frac{2\pi}{12} = 0.524$<br>$c = 4$, denn<br>$\text{Extremum}\left(\frac{17}{2} \sin\left(\frac{131}{250} x\right) + \frac{101}{10}, -4, -2\right)$<br>$\rightarrow \ (-3, 1.6)$<br>Daraus folgt, dass es ein Minimum gibt bei $x = -3$ das nach oben auf $x = 1$ verschoben werden muss. | 1 | 1<br>1 | 1<br>1 | |
| | An einem bestimmten Tag im März 2002 wurde die Niederschlagsmenge beobachtet. Die Niederschlagsmenge an diesem Tag kann durch die Funktion $R$ modelliert werden mit<br>$R(t) = 0.002t^3 - 0.064t^2 + 0.512t, 0 \leq t \leq 24$<br>wobei t die Zeit in Stunden ist und R(t) die Niederschlagsmenge in mm/h . | | | | |
| d) | **Beschreiben** Sie anhand einer kurzen Textbeschreibung diesen Tag in Bezug auf die Niederschlagsmenge. Ihre Antwort sollte sich auf die Zeiten mit dem meisten und dem geringsten Niederschlag konzentrieren. | | | | |
| | Es beginnt um Mitternacht zu regnen und der Niederschlag erreicht zwischen 5 und 6 Uhr einen Höhepunkt. Der Regen wird bis 16 Uhr weniger, dann hört er kurzzeitig auf. Am Nachmittag wird der Regen wieder stärker, bis er um Mitternacht ein Maximum erreicht. | | | 2 | 1 |

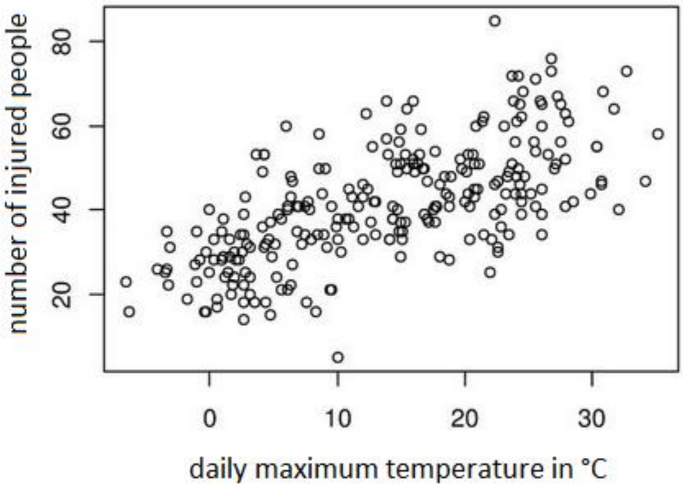
| | | | | | |
|---|---|---|---|---|---|
| |  | | | | |
| Ein leerer Glaszylinder wurde an diesem Tag nach draußen gestellt, um zu ermitteln, wie viel Regen gefallen war. | | | | | |
| e) | **Skizzieren** Sie den Graphen einer Funktion, der die Höhe des Wassers in einem Glaszylinder angibt, der zuvor im Freien stand und leer war. | | | | |
| |  Der Graph zeigt die Ableitungsfunktion von $R$. | 1 | 1 | 1 | |
| f) | **Berechnen** Sie die Gesamtmenge des Regens an diesem Tag in mm. | | | | |
| | $\int_0^{24} R(t) = 18.43$ An diesem Tag fielen etwa 18,4 mm Regen. | | 1 | 1 | |
| Das Jahr 2002 hatte im Großherzogtum Luxemburg 195 Regentage und 170 Tage ohne Regen. Es kann davon ausgegangen werden, dass alle Tage die gleiche Chance haben, ein Regentag zu sein. Ein Jahr später wollen Meteorologen untersuchen, ob es im Jahr 2003 mehr Regen gab. Leider sind einige Daten verloren gegangen, so dass sie nur eine kleine Stichprobe von 30 aufeinanderfolgenden Tagen genommen haben. | | | | | |
| g) | **Berechnen** Sie die Wahrscheinlichkeit, dass es an einem zufälligen Tag regnet, wenn wir annehmen, dass die Gesamtzahl der Regentage in beiden Jahren konstant bleibt und die Regentage gleichmäßig über das ganze Jahr verteilt sind. | | | | |
| | $p = \frac{195}{365} = 0.534$ | 1 | | | |
| h) | **Verwenden** Sie ein NHST-Verfahren, um herauszufinden, an wie vielen Tagen es regnen muss, damit die Meteorologen sagen können, dass es 2003 im Vergleich zu 2002 mehr geregnet hat, wenn das Signifikanzniveau bei 5% liegt. | | | | |
| | $H_0: p > 0.534$ <br> $H_1: p \leq 0.534$ | | 2 | | |




| k | P(X ≤ k) |
|---|---|
| 0 | 0 |
| 1 | 0 |
| 2 | 0 |
| 3 | 0 |
| 4 | 0 |
| 5 | 0 |
| 6 | 0.0002 |
| 7 | 0.0008 |
| 8 | 0.0027 |
| 9 | 0.0082 |
| 10 | 0.0214 |
| 11 | 0.049 |
| 12 | 0.0989 |
| 13 | 0.1781 |
| 14 | 0.2884 |
| 15 | 0.4232 |
| 16 | 0.5679 |
| 17 | 0.7046 |
| 18 | 0.8176 |
| 19 | 0.8995 |
| 20 | 0.951 |
| 21 | 0.9792 |

| | | | | | |
|---|---|---|---|---|---|
| | $k = 20$<br>Wenn es also an mehr als 20 Tagen geregnet hat, können die Wissenschaftler zu 95% sicher sein, dass es im Jahr 2003 mehr geregnet hat als im Jahr 2002. | | 1 | 1 | 1 |

Das folgende Diagramm zeigt die Höchsttemperatur und die Anzahl der Verletzten durch Verkehrsunfälle in Berlin im Langzeitvergleich.



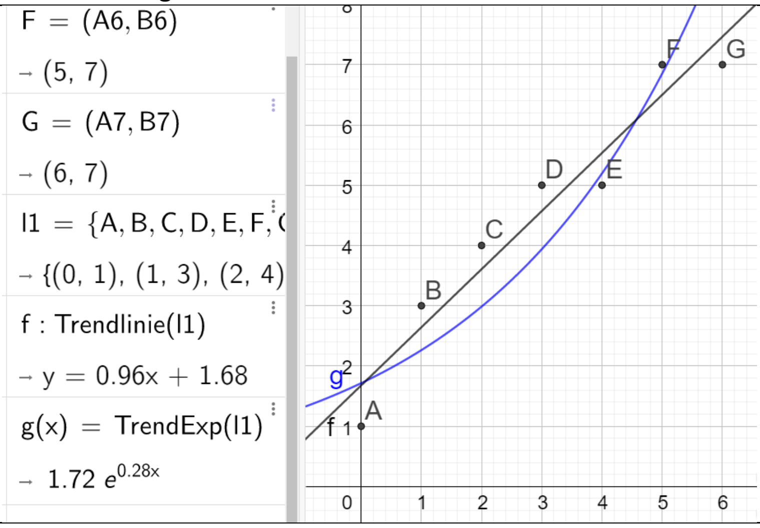

| | | | | | |
|---|---|---|---|---|---|
| i) | **Beschreiben** Sie die Korrelation zwischen den beiden Werten. | | | | |
| | Je höher die Temperatur, desto mehr verletzte Personen. | | | | 1 |
| j) | **Erklären** Sie, warum die Anzahl der Verletzten möglicherweise so mit der maximalen Temperatur korreliert. | | | | |
| | Wenn es warm ist, gehen mehr Menschen nach draußen, so dass die Wahrscheinlichkeit, in einen Autounfall verwickelt zu werden höher ist, als wenn man zu Hause bleibt. | | | | 1 |
| | **Gesamt = 25** | **6** | **8** | **7** | **4** |

| | Teil B2 | Punkte | | | |
|---|---|---|---|---|---|
| | | KV | M | P | I |
| | In einer Covid-19-Teststation wurden an einem bestimmten Tag 19 Personen mit Symptomen getestet und 6 von ihnen hatten ein positives Ergebnis. Am gleichen Tag wurden 87 Personen ohne Symptome getestet, von denen 85 negativ getestet wurden. | | | | |
| a) | **Zeigen** Sie, dass die Wahrscheinlichkeit, ein positives Ergebnis zu erhalten, davon abhängt, ob eine Person Symptome hat oder nicht. | | | | |
| | Es seien die folgenden Ereignisse: S: "die Person hat Symptome" P: "die Person ist positiv". $P(P) = \frac{8}{106}$ und $P(P\|S) = \frac{6}{19}$. $P(P) \neq P(P\|S)$ Die Wahrscheinlichkeit, ein positives Ergebnis zu erhalten, hängt also davon ab, ob eine Person Symptome hat oder nicht. | | 2 | | |
| | Zum Schutz persönlicher Daten sind die Testresultate mit einem Code versehen, der 2 Buchstaben (aus einem Alphabet mit 26 Buchstaben) und 4 Ziffern (0-9) enthält. Die gleichen Buchstaben und Ziffern dürfen mehrfach gewählt werden. | | | | |
| b) | **Berechnen Sie** die Gesamtzahl der verschiedenen Codes, die von diesem System erstellt werden können. | | | | |
| | Die Gesamtzahl der verschiedenen Codes ist $26^2 \cdot 10^4 = 6\,760\,000$. | 1 | 1 | | |
| | Nach einigen Monaten hat die Statistik gezeigt, dass 1,7 % der Personen ohne Symptome positiv getestet werden. Ein Unternehmen mit 20 Mitarbeitern (alle ohne Symptome) lässt alle testen. | | | | |
| c) | **Nennen Sie** zwei Annahmen, die getroffen werden müssen, um diese Situation mit einer Binomialverteilung zu modellieren. | | | | |
| | • Für die Binomialverteilung müssen alle Versuche unabhängig sein. Wir müssen also annehmen, dass die Mitarbeiter keine Möglichkeit haben, sich gegenseitig anzustecken.<br>• Wir wissen auch, dass manche Menschen (bestimmte Altersgruppen oder gesundheitliche Probleme) eine höhere Wahrscheinlichkeit haben, sich zu infizieren. Wir müssen also davon ausgehen, dass die Ansteckungsgefahr für alle Mitarbeiter gleich groß ist.<br>• Außerdem müssen wir davon ausgehen, dass die Grundgesamtheit der Statistik eine sehr große Anzahl von Menschen ist.<br>• Wir müssen davon ausgehen, dass die Tests zuverlässig sind.<br>• Wir müssen davon ausgehen, dass kein Mitarbeiter geimpft wurde.<br>• Wir müssen davon ausgehen, dass keiner der Mitarbeiter vorher Covid-19 hatte, so dass er jetzt immun sein könnte. | 1 | 1 | | |
| d) | **Berechnen** Sie die Wahrscheinlichkeit, dass mindestens einer der Mitarbeiter positiv getestet wird. | | | | |
| | Man rechnet die Wahrscheinlichkeit, dass keiner der Mitarbeiter positiv getestet wird: $P(X=0) = B_{200;0,017}(0) = \binom{200}{0} \cdot 0{,}017^0 \cdot 0{,}983^{200}$ Also gilt: $P(X \geq 1) = 1 - P(X=0) = 1 - 0{,}017^0 \cdot 0{,}983^{200}$ , Dies ergibt $P(X \geq 1) \approx 0{,}968$ | | 2 | 1 | |
| | Angenommen, die Situation kann durch eine Binomialverteilung modelliert werden, die durch die Formel $B(84; 0{,}02; k) = \binom{84}{k} \cdot 0{,}02^k \cdot 0{,}98^{84-k}$. | | | | |
| e) | **Interpretieren** Sie die Werte 84, 0,02 und 0,98 im gegebenen Zusammenhang. | | | | |
| | 84 ist die Gesamtzahl der Mitarbeiter. 0,02 ist die Wahrscheinlichkeit, dass ein Mitarbeiter positiv getestet wird und 0,98 ist das Komplement, d.h. die Wahrscheinlichkeit, dass ein Mitarbeiter nicht positiv getestet wird. | 1 | 2 | | |

Am 5. März 2020 ist ein aus Italien zurückgekehrter Mann die erste Person im Großherzogtum Luxemburg, die positiv auf COVID-19 getestet wurde. Daher ist dieser Tag in der Statistik als Tag 0 markiert. Die folgende Tabelle zeigt die Gesamtzahl der registrierten infizierten Personen in Luxemburg in den Tagen nach Auftreten des ersten Falles.

| Tag | 0 | 1 | 2 | 3 | 4 | 5 | 6 |
|---|---|---|---|---|---|---|---|
| Nummer | 1 | 3 | 4 | 5 | 5 | 7 | 7 |

f) **Zeichnen** Sie ein Streudiagramm dieser Werte zusammen mit einem linearen und einem exponentiellen Regressionsmodell.




| | | | |
|---|---|---|---|
| | 2 | 1 | |

g) **Geben** Sie die Gleichungen **an**, die die Regressionsgeraden in Frage f) beschreiben.

Linear: $f(x) = 0.96x + 1.68$
Exponential: $1.72 \cdot e^{0.28x}$

| | | | |
|---|---|---|---|
| 2 | | | |

h) **Erklären** Sie, warum es so schwierig war zu entscheiden, ob die Ausbreitung des Virus in diesem frühen Stadium am besten mit einem linearen oder einem exponentiellen Modell modelliert wird.

In einem frühen Stadium scheinen beide Regressionen gut zu den Daten zu passen. Die Ausreißer beeinträchtigen keines der Modelle kritisch. Es gibt zwar einen zusätzlichen Ausreißer für das Exponentialmodell, aber die Kurve geht durch zwei Punkte, während sie für das lineare Modell kaum zwei Punkte berührt.

| | | | |
|---|---|---|---|
| | 2 | | |

Nach weiteren sieben Tagen wurden andere Modelle erstellt, die bessere Vorhersagen machten, wobei t in Tagen angegeben ist:

$$A(t) = 1.35567 \cdot 1.46977^t$$
$$B(t) = 12.4396 \cdot t - 34.8571$$

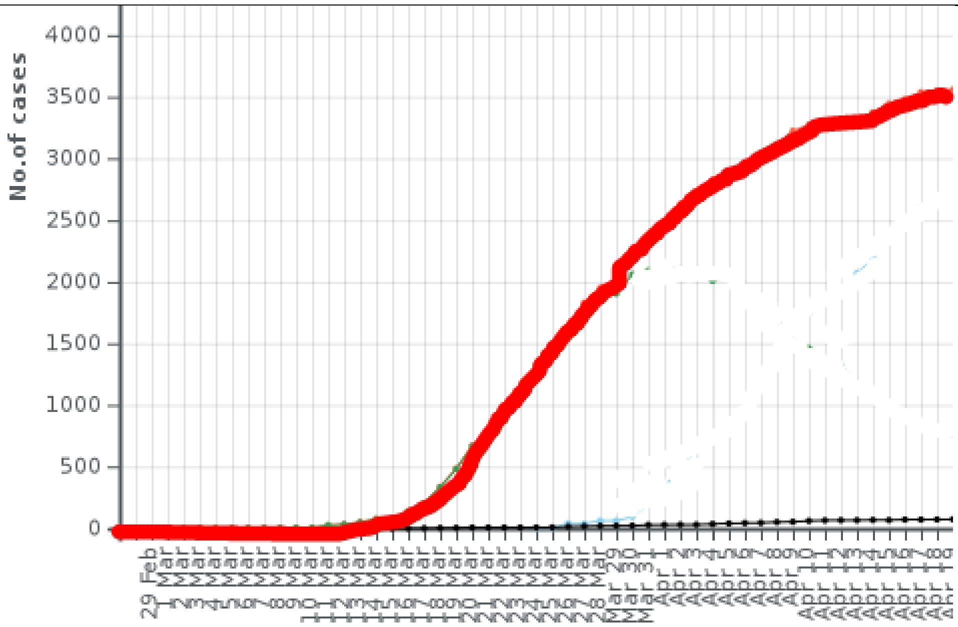
Am 16. Tag gab es im Großherzogtum Luxemburg 670 registrierte Fälle von COVID-19.

i) **Berechnen** Sie die vorhergesagte Anzahl der infizierten Personen an Tag 16 mit Modell A und Modell B und **vergleichen Sie** diese mit der wahren Anzahl. **Entscheiden** Sie, welches Modell für diese Situation offensichtlich besser funktioniert und **begründen** Sie Ihre Antwort.

$A(16) = 1{,}35567 \cdot 1{,}46977^{16}$ also gilt: $A(16) \approx 643$
$B(16) = 12.4396 \cdot t - 34{,}8571$ also gilt: $B(16) \approx 164$
Der vorhergesagte Wert für Modell A liegt nahe am beobachteten Wert. Der vorhergesagte Wert für Modell B ist viel kleiner als der beobachtete Wert. Angesichts der großen Abweichung zwischen den beiden vorhergesagten Werten ist Modell A besser für die Situation geeignet.

| | | | |
|---|---|---|---|
| | 1 | 1 | |

Das folgende Diagramm zeigt den Verlauf der Gesamtzahl der registrierten Infektionen für die ersten 4 Wochen im Großherzogtum Luxemburg.




<table>
<tr><td>j)</td><td colspan="5">Nennen Sie zwei mögliche Gründe, warum die Kurve im späteren Verlauf abflacht.</td></tr>
<tr><td></td><td>
• (Unter Verwendung des technologischen Hilfsmittels) können wir sehen, dass $\lim_{t\to\infty} C(t) = 3404$ was zeigt, dass die Kurve eine horizontale Tangente hat und daher abflacht, wenn $t$ zunimmt.<br>
• $\lim_{t\to\infty} C'(t) = 0$ , was zeigt, dass sich die Tangente an dem Graphen mit zunehmendem $t$ in eine horizontale Position bewegt und die Kurve daher abflacht.<br>
Es können auch nicht-mathematische Gründe angegeben werden:<br>
• Im Sommer gingen die Infektionsraten zurück, weil viele Menschen nach draußen gehen.<br>
• Es gab einen Lockdown, bei dem Geschäfte und Schulen geschlossen wurden, was einen Rückgang der Infektionsrate zur Folge hatte.<br>
• Die Leute trugen mehr Masken.<br>
Andere Gründe, die falsch sein könnten, aber trotzdem das Ergebnis erklären würden:<br>
• Menschen, die Covid-19 gehabt haben sind immun, so dass die Gesamtheit der Infektionen abnimmt.<br>
• Die Anzahl der Tests nahm ab.
</td><td></td><td></td><td>1</td><td>1</td></tr>
<tr><td colspan="6">Die Kurve kann durch die Funktion $C$ modelliert werden mit<br>$$C(t) = \frac{3404}{1+193\cdot e^{-0.233\cdot t}}$$</td></tr>
<tr><td>k)</td><td colspan="5">Ermitteln Sie den Tag mit der höchsten Infektionsrate durch Berechnung.</td></tr>
<tr><td></td><td>Wir müssen den Maximalwert für die Ableitung auf dem Intervall [0; 28] finden.<br>Mit dem technologischen Hilfsmittel finden wir, dass der maximale Wert für $C'$ erreicht wird bei $t = 22{,}59$ , d. h., am 23. Tag erreicht die Infektionsrate ihr Maximum.</td><td>1</td><td>1</td><td></td><td></td></tr>
<tr><td></td><td>Gesamt = 25</td><td>7</td><td>13</td><td>4</td><td>1</td></tr>
</table>