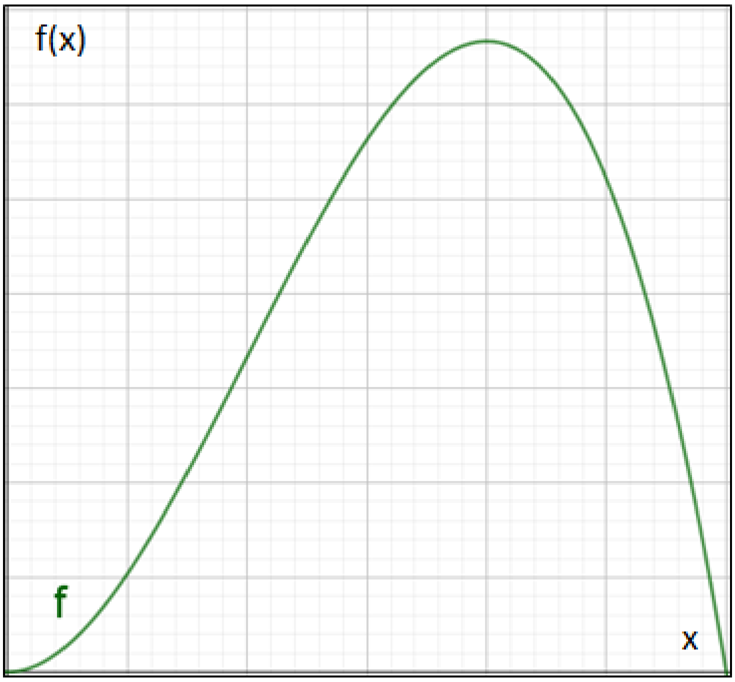
Part A Answers

	Part A	Points			
		KC	M	PS	I
1	A little hill on a playground can be modelled by a function f with $f(x) = -\frac{1}{3}x^3 + x^2$, for $x > 0$ where x is the distance in m and $f(x)$ is the height in m. The picture shows the graph f of this function.  **Determine** the height of this hill.				
	$f'(x) = -x^2 + 2x$	1	1		
	$f'(x) = 0$	1			
	$-x^2 + 2x = 0$				
	$x \cdot (-x + 2) = 0$				
	$x = 0 \vee -x + 2 = 0$				
	$x = 0 \vee x = 2$		1		
	$f(2) = -\frac{8}{3} + 4 = -\frac{8}{3} + \frac{12}{3} = \frac{4}{3} \approx 1.33$				
	The height is about 1.33 m.		1		
2	After some complaints about the lunch in the canteen, the manager claims that at most only 20% of all 2,500 pupils are not satisfied with the lunch. The pupils committee thinks that it is more than 20% of the pupils. So, they ask a group of 40 randomly chosen pupils for their opinion. a) **Explain** whether a left or a right sided test should be used to verify this hypothesis. Reason your answer. b) **State** the null hypothesis H_0 that could be used for a NHST test and give the alternative hypothesis H_1. c) **Determine** the critical value k with the help of the following table if the significance level is set at 5% and interpret this value.				

k	8	9	10	11	12	13	14	15
$P(X \geq k)$	0.563	0.407	0.268	0.161	0.088	0.043	0.019	0.008

a) A right sided test is appropriate, because the pupils want to show, that p is higher.	1	1		
b) $H_o: p \leq 0.2$ and $H_1: p > 0.2$	1			
c) $k = 13$, because $P(X = 12) > 0.05$ and $P(X = 13) < 0.05$ That means, if 13 of more students are not satisfied, the manager is proved to be wrong.			1	1

3 A small supermarket chain employs 900 people, 10 of them work in the management, but only one of the managers is female. The other 809 women work in the shops.
Show, that it depends on the sex, if you get a position in the management of this company.

The given information can be put into a contingency table to reconstruct the missing values.

M: management
F: female

	F	$\bar{F}$	
M	1	9	10
$\bar{M}$	809	81	890
	810	90	900

	1	1	

Dependency of the two events can be shown by using the formula.

$$P(F \cap M) = \frac{1}{900}$$

$$P(F) \cdot P(M) = \frac{810}{900} \cdot \frac{10}{900} = \frac{9}{10} \cdot \frac{1}{90} = \frac{9}{900} = \frac{1}{100}$$

1		1	
	1		

$$P(F \cap M) \neq P(F) \cdot P(M)$$

4 A couple needs a negative covid test to visit friends abroad. It is known that 20 % of the tests show a negative result, although the person could be infected (false negative result). The probability of a false positive result is close to zero. It can be assumed that if one of them is infected, that the other one is also infected.
Explain, why this situation is a Bernoulli process and **show**, that the probability of a false negative result drops down to 4 % when both get tested.

Both tests are independent. There are only two outcomes, the test can be right or wrong and the probability of a wrong result is the same in each test. $P(X = 2) = B(2; 0{,}2; 2) = 0{,}2^2 = 0{,}04 = 4\%$	2	3		

5 In the diagram shown below, the number of new COVID-19 cases in Germany is shown over a period of 3 weeks in October 2020. To predict the numbers in the future, two basic types of mathematical models can be combined.

	Give the names of these types of models and **reason** your answer. **Predict** a date in the future when another maximum will be reached, when the numbers follow the models.				
	Since the numbers are oscillating with a period of 7 days, a periodic model can be used on a small-scale level. However, the numbers are altogether increasing on a large-scale level. To model this development, an exponential model could be used, because neither maxima nor minima fit a straight line. The next maximum will be reached on 24. October.	1	1	2	1
6	The number of bacteria in a petri dish is investigated in a laboratory. It turns out, that under certain conditions, the growth can be modelled by the function $$N(t) = 10\ 000 \cdot e^{\ln(1{,}03)\cdot t},$$ where $N(t)$ is the number of bacteria after t days. a) **Give** the number of bacteria at the beginning and the growth rate in percent. b) **Calculate** the number of bacteria after the first day. c) **Explain**, why this model cannot be used on a very large time scale.				
	a) The experiment starts with 10 000 bacteria. Per day, the number increases by 3%. b) $N(1) = 10\ 000 \cdot 1{,}03 = 10\ 000 + 300 = 10\ 300$ c) At some stage, the petri dish is full of bacteria and there is no more space or nutrition, so the growth will stop.	2	2		1
7	**Indicate** if the statement is true or false and **reason** your answer. Note that the points are only given if answer and reason are correct. a. If the temperature $T(x)$ is constantly increasing, then $T'(x) > 0$. b. All periodic models can be modelled by a sine function. c. There are 9 different possibilities for 3 pupils to stand next to each other. d. When some dice is rolled once, the expected value is 3.5. e. If 10 people are chosen out of a very large group, the number of females can be modelled by a binomial distribution, although a person cannot be chosen more than once.				
	a) True, because the if the temperature is increasing, the rate of chance is positive, and the rate of change is given by the derivative.		1		

	b) False, because the red light of a traffic light be a periodic model. But the light is either on or off with no stages in between, so it cannot be modelled by a sine function.	1			
	c) False, because this example is without repetition. So, in fact there are 6 possibilities.			1	
	d) True, because $E(X) = (1+2+3+4+5+6):6 = 3.5$		1		
	e) True, because the change in the probability is rather small, when we have a large group. Beside a constant probability, there can only be two options to choose from (male and female) and it must be a random choice.		1		
8	The daylength $L(t)$ in hours on a certain location was recorded over one year. It can be modelled by the function $$L(t) = 4 \cdot \sin\left(\frac{2\pi}{365}x\right) + 12,$$ where t is the time in days. **Interpret** the outcome of $\int_0^{365} L(t)dt$ and **explain**, why the result is equal to $12 \cdot 365 = 4380$.				
	The total amount of daylight hours in one year can be calculated by the integral of the function between 0 and 365: $$\int_0^{365} L(t).dt$$				1
	The sine function is symmetrical, so the area above the x-axis has the same value as the area below the x-axis on the interval of one period. So, the integral would be zero. Since L(t) is shifted by 12 units and we calculate the area under the graph for 365 days, the result is $12 \cdot 365 = 4380$. One could also argue that the benefit of long days in summer is always matched by long periods of darkness in winter; thus, the total amount can be expressed by 12·365=4.380 hours.		1	3	
9	a) **Interpret** what is meant by expected value of a random variable. b) X is a random variable following a normal distribution with expected value μ and standard deviation σ. **Give** a probability taking into account these two characteristic values μ and σ. c) A continuous random variable Y defined over $\mathbb{R}$ is such that $P(a \le y \le b) = \int_a^b f(z)dz$. **Explain** why $\int_{-\infty}^{+\infty} f(z)dz = 1$.				
	a) The expected value of a random variable is the average of the values that this random variable takes weighted by their probabilities.	2			
	b) $P(\mu - \sigma \le Y \le \mu + \sigma) \approx 0{,}68$ or $P(\mu - 2\sigma \le Y \le \mu + 2\sigma) \approx 0{,}95$ or $P(\mu - 3\sigma \le Y \le \mu + 3\sigma) \approx 0{,}997$	1			
				1	1

	c) The integral $\int_{-\infty}^{+\infty} f(z)dz$ corresponds to $P(-\infty < Z < +\infty)$, the probability for Y to take all possible real values. Since Y denotes a random variable defined on $\mathbb{R}$, this probability is therefore that of all possible events, hence the value is 1.				
10	A new machine recognises doping in blood. Let there be the following two events • P: The test is positive. • D: The sportsman was doped. After some test runs it was found out, that out of 100 blood samples with doping, the machine recognises it in 90 cases. However, it also gives false alarm in 5% of the cases, when the sample was clean. It can be assumed, that every 10th sportsman at a certain event is doped. We want to find out the probability that a sportsman was indeed doped when the test is positive. a) **Present** all necessary information in the correct mathematical notation. b) **Use** an appropriate method to determine the probability that a sportsman was doped, given that the test was positive.				
	a) $P(P\|D) = 0.9$ $P(P\|\overline{D}) = 0.05$ $P(D) = 0.1$ b) For example Bayes Theorem: $P(D\|P) = \frac{P(P\|D) \cdot P(D)}{P(P)} = \frac{P(P\|D) \cdot P(D)}{P(P\|D) \cdot P(D) + P(P\|\overline{D}) \cdot P(\overline{D})}$ $= \frac{0.9 \cdot 0.1}{0.9 \cdot 0.1 + 0.05 \cdot 0.9} = \frac{0.1}{0.15} = \frac{10}{15} = \frac{2}{3}$ A tree diagram could also be used.	1	1 2	1	
	Total = 50	**14**	**21**	**9**	**6**