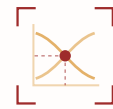


FSP Training 07

Preparation for Mock Exam 08



Instructions

This training page covers all of Section 07 and prepares you for Mock Exam 08. Set A trains setting up: you translate a text into mathematics and check your translation. Set B is a practice mock in the format of Mock Exam 08: 180 minutes plus 5 minutes reading time, 100 points, and all 3 problems must be completed. Write down every step as you would in the exam. The points show what each part is worth.

Round probabilities to four decimal places (or give them in percent with two decimal places) and keep intermediate results unrounded. Round descriptive statistics such as the mean or the standard deviation to two decimal places. A minimum number n is always rounded up: check $n - 1$ and n .

In the real FSP, decimals are written with a comma (1,25 means 1.25).

Set A: Set Up & Check (40 minutes)

From Text to Mathematics (xx)

Translate each statement into mathematics. Then answer the check question: it tells you whether your translation works.

1. A salad bar prepares 60 salad bowls every morning. On 8 randomly chosen days, the numbers of bowls left unsold at closing time were 10, 4, 13, 7, 14, 10, 3 and 11.

Compute the mean and the sample standard deviation of the daily number of unsold bowls. **Write** the probability that a bowl stays unsold as a relative frequency and **compute** it.

Check: A classmate writes $P(\text{unsold}) = \frac{72}{60} = 1.2$. **Explain** why this number cannot be a probability and **give** the correct base.

2. In a public library, 70 % of the visitors borrow a book (B), 20 % use a study place (L), and 14 % do both. **Decide** with $P(B \cap L) = P(B) \cdot P(L)$ whether B and L are independent. **Write** “a visitor does neither” as a set and **compute** its probability.

Check: A classmate computes the share of the visitors who do neither as $1 - 0.70 - 0.20 = 0.10$. **Draw** a Venn diagram with the probabilities of its four regions and **explain** the mistake.

3. A school receives 13 new laptops. 4 of them have a faulty keyboard, but nobody knows which. The IT teacher tests 3 of the laptops, chosen at random. **Write** the probability that exactly one tested laptop has a faulty keyboard as a quotient of counts and **compute** it. **Compute** the probability that at least one tested laptop has a faulty keyboard.

Check: A classmate draws a tree and computes $\frac{4}{13} \cdot \frac{9}{12} \cdot \frac{8}{11}$ for “exactly one”. **Compute** this product and **explain** why it is smaller than your result.

4. At a car-sharing company, 45 % of the rentals are made by new customers (N), the rest by regular customers. 2 % of the rentals by regular customers end with a parking fine (F). In total, 5.6 % of all rentals end with a parking fine. **Set up** the law of total probability with the unknown $x = P(F | N)$ and **show** that $x = 0.1$.

Check: **Put** $x = 0.1$ into the tree and **add** the two paths that end in F . A classmate computes $x = 0.056 - 0.02 = 0.036$. **Explain** the mistake.

5. A heating company services gas boilers at a fixed price of 149 EUR per service. From experience, a service costs the company 100 EUR with probability 0.7, 180 EUR with probability 0.25 (a small spare part is needed) and 400 EUR with probability 0.05 (a larger repair that the fixed price includes). X is the company’s cost of one service. **Write** the expected cost as a formula and **compute** it. **Decide** whether the fixed price covers the costs on average.

Check: A classmate takes the mean of the three costs, $\frac{100+180+400}{3} \approx 226.67$ EUR, and says that the company loses money on every service on average. **Explain** the mistake.

6. A sawmill scans every log for hidden metal (nails, old fence wire) before cutting it, because metal destroys the saw blades. 1 % of the logs contain metal (M). The scanner raises an alarm (+) for 97 % of the logs with metal and for 4 % of the logs without metal. **Write** the three rates in probability notation and **name** them. **Compute** the probability that a log with an alarm does indeed contain metal, and the probability that a log contains metal but is incorrectly identified as free of metal.

Check: **Complete** a table of counts for 10,000 logs and **compute** both probabilities again from the table. The scanner’s supplier says: “Our scanner is right in 97 % of its alarms.” **Decide** whether this is true.

7. A high ropes course (a climbing park in the trees) had 1,200 visitors last month. 480 of them came with a school class (S). Of the visitors with a school class, 15 % booked the lunch package (F). In total, 300 visitors booked the lunch package. **Complete** the contingency table of counts. **Write** the expected count of the field $S \cap F$ if booking the lunch package were independent of coming with a school class, **compute** it and **compare** it with the observed count.

Check: A classmate computes the expected count as $0.4 \cdot 0.15 \cdot 1200 = 72$ and says: “Observed and expected count agree, so the two properties are independent.”

Explain the mistake.

8. A removal company carries out 15 moves per week. In each move, some furniture gets damaged with probability 15 %, independently of the other moves. X is the number of moves with damage in one week. **Write** “more than 2 but fewer than 6 moves with damage” as a probability with the correct endpoints and as a sum of binomial terms, and **compute** it. **Write** “at least 3 moves with damage” with the complement and **compute** it.

Check: A classmate writes the first event as $P(2 \leq X \leq 6)$. **List** the values of X in the classmate's event and in the event from the text.

One Paragraph, All Conditions (xx)

A company maintains the lifts in office buildings. Once a day, a remote sensor checks every lift and reports a possible fault (+). On a given day, 6 % of the lifts have a fault (F). The sensor reports 80 % of the lifts with a fault, and it also reports 3 % of the lifts without a fault. After every report, a technician drives to the lift. The manager wants to know the probability that a technician who drives out finds a real fault, the probability that a lift has a fault but the sensor does not report it, and the number of trips to expect on one day for the 250 lifts of the company. The manager also asks for the probability that exactly 2 of the 12 lifts in one office tower are reported on the same day, and for the smallest number of lifts that must be checked so that, with a likelihood (probability) of at least 99 %, at least one lift with a fault is reported.

List every given value and every question in this paragraph as a mathematical statement, with the course name of each rate. **Name** the rule or the distribution you would use. Do not solve anything.

Check: An assistant writes the last question as $1 - 0.2^n \geq 0.99$. **Explain** what went wrong.

FSP Words So Far

The FSP problems use these words from Sections 01 to 07. Learn what each one means and how to write it in mathematics.

Word	Meaning	Mathematics
quantity unit (QU)	a fixed number of items, e.g. 1 QU = 100 items	the output x is counted in QU
monetary unit (MU)	a fixed amount of money, e.g. 1 MU = 100 EUR	results in MU, convert to EUR at the end
fixed costs	costs that arise even if nothing is produced	the constant term of $C(x)$
variable costs	costs that grow with the output	$C(x) - C(0)$ (for linear costs: output times costs per QU)
total costs	fixed costs + variable costs	$C(x)$
revenue	quantity times price	$R(x) = x \cdot \text{price}$
profit	revenue – total costs	$P(x) = R(x) - C(x)$
break-even point	the output at which the profit changes from negative (a loss) to positive	$P(x) = 0$, the smaller zero

Word	Meaning	Mathematics
profit limit (upper break-even point)	the larger output with zero profit: above it, the profit turns negative again	$P(x) = 0$, the larger zero
profit zone	all outputs with a positive profit	$P(x) > 0$: from the break-even point to the profit limit (or the capacity)
capacity	the maximum possible output	an upper bound: $x \leq$ capacity
growth factor	the factor by which a quantity is multiplied each period	q in (starting value) $\cdot q^t$
growth rate	the change per period in percent	$q = 1 + r$ with r as a decimal, a loss of 4 % gives $q = 0.96$
doubling time	the time after which a quantity has doubled	$q^T = 2$, continuous: $e^{kT} = 2$, so $T = \frac{\ln 2}{k}$
half-life (halving time)	the time after which a quantity has halved	$q^T = 0.5$, continuous: $e^{-kT} = 0.5$, so $T = \frac{\ln 2}{k}$
“deviates from a by at most d ” (tolerance)	the distance from a is at most d	$ x - a \leq d \Leftrightarrow a - d \leq x \leq a + d$
“share of A among B ”	the part of B that also belongs to A : the base is B	$\frac{\text{number in } A \cap B}{\text{number in } B}$
“neither A nor B ”	not in A and not in B , i.e. outside both circles	$\overline{A \cup B}$
contingency table	a table of counts for two properties: the inner fields are the regions of the Venn diagram, the margins are the totals	inner fields $A \cap B$, $A \cap \overline{B}$, $\overline{A} \cap B$, $\overline{A} \cap \overline{B}$. Joint probability: field divided by the grand total, conditional probability: field divided by the total of the condition's row or column
gross price, net price	gross price = net price + VAT	with 19 % VAT: gross = $1.19 \cdot$ net
system of linear equations	several linear equations that must hold at the same time	solve by substitution or elimination, check the solution in every equation
no solution, infinitely many solutions	the conditions contradict each other, or one condition only repeats another	elimination leaves $0 = d$ with $d \neq 0$, or $0 = 0$

Word	Meaning	Mathematics
demand, supply	the quantity buyers want to buy and the quantity sellers offer at a given price	as terms in the price p , e.g. $Q_d(p)$, $Q_s(p)$, or as prices $p_D(x)$, $p_S(x)$ for the quantity x
market equilibrium	the quantity demanded equals the quantity supplied	demand = supply: $Q_d(p) = Q_s(p)$, or $p_D(x) = p_S(x)$. With demand and supply as prices: $D(q) = S(q)$ gives q^* , and $p^* = D(q^*)$
price-demand function	gives the price at which x QU can be sold	$p(x)$, revenue $R(x) = x \cdot p(x)$
average cost	total costs divided by the number of units	$\frac{C(x)}{x}$ with $x \neq 0$
zero form	all terms on one side, 0 on the other side	$ax^2 + bx + c = 0$, then factor or use the quadratic formula
discriminant	decides how many real solutions a quadratic equation has	$\Delta = b^2 - 4ac$: > 0 two, $= 0$ one, < 0 none
biquadratic equation	an equation with only x^4 , x^2 and a constant	substitute $u = x^2$ in $ax^4 + bx^2 + c$, solve for u , then $x = \pm\sqrt{u}$ for $u \geq 0$
domain restrictions (of an equation)	the conditions the unknown must meet so that every term is defined	denominators $\neq 0$, under a root ≥ 0 , inside a logarithm > 0
extraneous solution	a candidate that solves the squared equation but not the original one	check every candidate in the original equation
function	assigns to every input exactly one output	$f(x)$, its graph passes the vertical line test
domain (of a function)	all inputs for which the function is defined, in a context only the sensible ones	e.g. $[0; \text{capacity}]$
range	all outputs that the function takes on its domain	if the function grows or falls: between the values at the two ends, for a parabola also check the vertex
maximum price	the price at which nobody buys any more	$Q_d(p) = 0$, or c in $p = c - dQ$

Word	Meaning	Mathematics
market saturation	the quantity demanded at a price of 0	a in $Q_d = a - bp$, or the zero of $p(x)$
shortage, surplus	at a price below the equilibrium price, demand exceeds supply (shortage), above it, supply exceeds demand (surplus)	$Q_d(p) > Q_s(p)$, or $Q_d(p) < Q_s(p)$
contribution margin	price minus variable costs per unit: the part of the price that covers the fixed costs	break-even quantity = $\frac{\text{fixed costs}}{\text{contribution margin}}$
vertex	the highest or lowest point of a parabola	$x_v = -\frac{b}{2a}$, or (h, k) in $a(x - h)^2 + k$
profit-maximizing output	the output with the highest profit	x_v of $P(x)$, a maximum since $a < 0$, and the price $p(x_v)$. If the capacity lies below x_v : the capacity and its price
shift, stretch, compression	moving or scaling a graph	$a \cdot f(b(x - h)) + k$: first b , then h , then a , then k
mirror image at the y -axis	the graph flipped from right to left	$f(-x)$
mirror image at the x -axis	the graph flipped upside down	$-f(x)$
composite function	first the inner function, then the outer function	$(f \circ g)(x) = f(g(x))$
inverse function	reverses a function: gives the input that belongs to an output	solve $y = f(x)$ for x , e.g. $x(p)$ from $p(x)$
polynomial of third degree	a function with powers of x up to x^3	$ax^3 + bx^2 + cx + d$ with $a \neq 0$: four unknowns, so four conditions
degree, leading coefficient	the highest power of x , and the number in front of it	together they decide the end behavior
zero of multiplicity m	the factor $(x - c)$ appears m times in the factored form	odd m : the graph crosses the x -axis at c , even m : it touches it
polynomial long division	divides a polynomial by $(x - c)$ once c is a known zero	$f(x) = (x - c) \cdot q(x)$, the remainder must be 0
asymptotic behavior (end behavior)	what happens to $f(x)$ for $x \rightarrow \infty$ and $x \rightarrow -\infty$	polynomial: degree and sign of the leading coefficient.

Word	Meaning	Mathematics
		Rational function: its asymptotes
intercepts	the points where the graph meets the axes	y -intercept $(0, f(0))$, x -intercepts from $f(x) = 0$
rational function	a quotient of two polynomials	$\frac{p(x)}{q(x)}$ with the domain $q(x) \neq 0$
vertical asymptote	a vertical line near which the values grow without bound	a zero of the denominator that does not cancel
hole	a single missing point of the graph	a factor that cancels, its y -value comes from the simplified term
horizontal asymptote	a horizontal line that the graph approaches for $x \rightarrow \pm\infty$	$y = c$ with $f(x) \rightarrow c$
oblique asymptote	a slanted line that the graph approaches for $x \rightarrow \pm\infty$	the linear part of the long division, if the numerator's degree is one higher
grows less than proportionally	grows, but twice the input gives less than twice the output	$a \cdot x^b$ with $0 < b < 1$
continuous growth rate	the rate k of a model with the base e	$N_0 \cdot e^{kt}$, growth factor per period e^k
saturation level	a level that a quantity approaches but never exceeds	a horizontal asymptote $y = L$, e.g. of $L - a \cdot e^{-kt}$ or $\frac{L}{1 + Ae^{-kt}}$
behavior for $t \rightarrow \infty$	the value a quantity approaches in the long run	$\lim_{t \rightarrow \infty} N(t)$, often with $e^{-kt} \rightarrow 0$
radian	an angle measure in which a full turn is 2π	$180^\circ = \pi$, arc length $s = r\theta$
amplitude	half the distance between the highest and the lowest value	$ A = \frac{\max - \min}{2}$
midline	the horizontal line halfway between the highest and the lowest value	$y = D = \frac{\max + \min}{2}$
period	the length of one full cycle	$B = \frac{2\pi}{\text{period}}$
phase shift	the horizontal shift of a sine or cosine curve	C in $A \sin(B(t - C)) + D$ or $A \cos(B(t - C)) + D$
mirror image at the midline	the graph flipped upside down about its midline $y = D$	$2D - f(t)$

Word	Meaning	Mathematics
limit	the value that $f(x)$ approaches when x approaches a number a , or $\pm\infty$	$\lim_{x \rightarrow a} f(x)$, one-sided: $x \rightarrow a^-$ (from the left), $x \rightarrow a^+$ (from the right)
continuous (at a)	the graph has no jump and no gap at a	$f(a)$ is defined, $\lim_{x \rightarrow a} f(x)$ exists, and both are equal
removable discontinuity	a gap that a single value could close, e.g. a hole	the limit exists, but $f(a)$ is not defined or differs from it
derivative	the instantaneous rate of change: the slope of the tangent	$f'(x) = \lim_{h \rightarrow 0} \frac{f(x+h) - f(x)}{h}$
secant, slope of the secant	the line through two points of the graph, and its slope (the average rate of change)	$\frac{f(b) - f(a)}{b - a}$ on $[a; b]$
tangent, tangential gradient	the line that touches the graph at a point, and its slope	$y = f'(a)(x - a) + f(a)$ with the slope $f'(a)$
angle of intersection (with the x -axis)	the angle at which a line with slope m crosses the x -axis	$\alpha = \arctan(m)$, with the indicator D for degrees
marginal cost, marginal revenue, marginal profit	about the change of the costs, the revenue or the profit caused by one more unit	$C'(x), R'(x), P'(x) = R'(x) - C'(x)$
stationary point (critical point)	a point of the graph with a horizontal tangent	$f'(x_0) = 0$
nature of a stationary point	whether it is a local maximum, a local minimum or a saddle point	sign change of f' at x_0 , or the sign of $f''(x_0)$
local maximum, local minimum	a point that is higher (lower) than all points nearby	$f'(x_0) = 0$ and $f''(x_0) < 0$ (maximum) or $f''(x_0) > 0$ (minimum)
saddle point	a stationary point at which f' does not change its sign: the graph levels off and keeps its direction	$f'(x_0) = 0$ without a sign change, e.g. $f''(x_0) = 0$ and $f'''(x_0) \neq 0$
necessary condition, sufficient condition	a condition every candidate must meet, and a condition that proves the kind of point ("prove it really is a maximum")	necessary: $f'(x_0) = 0$, sufficient: e.g. $f''(x_0) < 0$ for a maximum

Word	Meaning	Mathematics
global maximum, global minimum (on an interval)	the highest or lowest value on the whole interval	compare the local extrema with the values at the ends of the interval
concave up, concave down	the graph curves upward like a cup, or downward like a cap	$f''(x) > 0$, or $f''(x) < 0$
inflection point	the point at which the curvature changes sign: the graph is steepest or flattest there	$f''(x_0) = 0$ and $f'''(x_0) \neq 0$, or a sign change of f''
“the curvature changes sign”	the graph changes from concave down to concave up, or back	an inflection point: $f''(x_0) = 0$. For a cubic cost function, the marginal costs are lowest there
inflectional tangent	the tangent at an inflection point: it shows the fastest growth or decline	$y = f'(a)(x - a) + f(a)$ with $f''(a) = 0$
unit costs, lowest unit costs	the total costs per unit, and the output at which they are lowest	$k(x) = \frac{C(x)}{x}$, minimum from $k'(x) = 0$ and $k''(x) > 0$
long-term price floor	the lowest unit costs: below this price, the company cannot cover its total costs in the long run	the minimum value of $k(x) = \frac{C(x)}{x}$
variable costs per unit	the variable costs divided by the output	$k_v(x) = \frac{C(x) - C(0)}{x}$
short-term lower limit price	the lowest variable costs per unit: below this price, the company cannot even cover its variable costs	the minimum value of $k_v(x)$
function family (German: <i>Funktionsschar</i>)	functions with the same term that differ only in a parameter	$f_a(x)$, e.g. $f_a(x) = x^2 - ax + 1$
parameter	a letter that is fixed within one function of the family but changes from one function to the next	solve a condition for it, e.g. $f'_a(2) = 0$
antiderivative (German: <i>Stammfunktion</i>)	a function whose derivative is f	$F'(x) = f(x)$
constant of integration	the free constant: two antiderivatives of f differ only by a constant	$F(x) + C$, fixed by one condition, e.g. $C(0) = \text{fixed costs}$ or $P(0) = -\text{fixed costs}$

Word	Meaning	Mathematics
indefinite integral	the family of all antiderivatives	$\int f(x) dx = F(x) + C$
definite integral	a number: the signed area between the graph and the x -axis from a to b	$\int_a^b f(x) dx = F(b) - F(a)$ (Fundamental Theorem of Calculus)
lower limit, upper limit	the start a and the end b of the interval in $\int_a^b f(x) dx$	“in the second month”: from $t = 1$ to $t = 2$
signed area	area above the x -axis counts positive, area below it negative	$\int_a^b f(x) dx$, it can be negative or 0
area (between a graph and the x -axis)	always a positive number	split at the zeros and add the absolute values of the integrals of the pieces
net change	the total change of a quantity, computed from its rate of change	$F(b) - F(a) = \int_a^b F'(x) dx$, e.g. extra costs $\int_a^b C'(x) dx$
“total ... within the first ... months”	everything that has come together from the start up to month b	$\int_0^b (\text{rate}) dt$
area enclosed by two graphs	the region between the graphs from one intersection point to the next	$\int_a^b [f(x) - g(x)] dx$ with f on top, limits from $f(x) = g(x)$, split where the graphs cross
“both regions have equal size”	two regions have the same area (not: the integral over both is 0)	compute both areas separately and compare them
consumer surplus	what the buyers gain: they would have paid more than the market price	$CS = \int_0^{q^*} [D(q) - p^*] dq$
producer surplus	what the sellers gain: they would have sold for less than the market price	$PS = \int_0^{q^*} [p^* - S(q)] dq$
total surplus	consumer surplus + producer surplus	$\int_0^{q^*} [D(q) - S(q)] dq$, the area between demand and supply
minimum price	a price set by the state: no unit may be sold for less	above p^* : only the quantity demanded at this price is traded, and the surpluses end there. Below p^* : no effect
average value	the constant height that gives the same area (the same total) over the interval	$f_{avg} = \frac{1}{b-a} \int_a^b f(x) dx$, “average monthly ... within

Word	Meaning	Mathematics
		the first b months": $\frac{1}{b} \int_0^b f(t) dt$
"prove that F is an antiderivative of f "	show it by differentiating, no integrating needed	compute $F'(x)$, e.g. with the product and chain rule, and show $F'(x) = f(x)$
integration by parts	the integration rule that reverses the product rule	$\int u dv = uv - \int v du$, choose u with LIATE
mean	the sum of all values divided by their number	$\bar{x} = \frac{1}{n} \sum x_i$
median	the middle value of the sorted data	odd n : the middle value, even n : the mean of the two middle values
quartiles	Q_1 and Q_3 : the medians of the lower and the upper half of the sorted data	interquartile range IQR = $Q_3 - Q_1$, the middle 50 % of the data. With an odd n , leave the median out of both halves
box plot	a picture of the five-number summary	a box from Q_1 to Q_3 with a line at the median, whiskers to the smallest and the largest value that are not outliers
outlier	a value far away from the rest of the data	a value outside the fences, i.e. outside $[Q_1 - 1.5 \cdot \text{IQR}; Q_3 + 1.5 \cdot \text{IQR}]$
sample standard deviation	the typical distance of the values of a sample from their mean	$s = \sqrt{\frac{\sum (x_i - \bar{x})^2}{n-1}}$, for a whole population divide by N
sample space	the set of all possible outcomes of a random experiment	S , e.g. $S = \{1, 2, \dots, 6\}$ for one die
event	a set of outcomes	$A \subseteq S$, with the probability $P(A)$ between 0 and 1
complement (German: <i>Gegenwahrscheinlichkeit</i>)	"not A ", everything outside A	$P(\bar{A}) = 1 - P(A)$, e.g. "at least one" = $1 - \text{"none"}$
mutually exclusive	the events cannot happen together	$A \cap B = \emptyset$, so $P(A \cup B) = P(A) + P(B)$
independent	one event does not change the probability of the other	$P(A \cap B) = P(A) \cdot P(B)$, or $P(A B) = P(A)$, or

Word	Meaning	Mathematics
		observed count = expected count
“of the ...”, “among the ...”, “given that ...”	names the group that a percentage refers to: the condition	behind the bar: “of the members, 30 % redeem” is $P(R M) = 0.3$
tree diagram	the stages of a random experiment, with a probability on every branch	the branches that leave one point add up to 1, the second stage carries conditional probabilities
path probability	the probability of one path through the tree	multiply along the path (multiplication rule), all path probabilities add up to 1
law of total probability	the probability of an event that several paths lead to	add these paths: $P(B) = P(A) \cdot P(B A) + P(\bar{A}) \cdot P(B \bar{A})$
Bayes' theorem	reverses a conditional probability	$P(A B) = \frac{P(A) \cdot P(B A)}{P(B)}$, with $P(B)$ from the law of total probability
base rate (prevalence)	the share of all items (or people) that have the property, before any test	$P(D)$
true positive rate (sensitivity)	the share of the items with the property that test positive	$P(+ D)$
true negative rate (specificity)	the share of the items without the property that test negative	$P(- \bar{D})$
false positive, false negative	a positive result without the property, a negative result with the property	$\bar{D} \cap +$ and $D \cap -$, with the rates $P(+ \bar{D}) = 1 - \text{true negative rate}$ and $P(- D) = 1 - \text{true positive rate}$
“... with a positive test result does indeed ...”	among the positive results, the ones that are right	the positive predictive value $P(D +)$, the base is all positive results
“... has ... but is incorrectly identified as ...”	an item has the property, but the test says it has not, seen among all items	the joint probability $P(D \cap -) = P(D) \cdot P(- D)$, one path of the tree
positive predictive value, negative predictive value	the share of the positive results that are right, and the share of the negative results that are right	$P(D +)$ and $P(\bar{D} -)$

Word	Meaning	Mathematics
“Display the fractions in a contingency table. Fill in the blank fields.”	write every given share into its field or total, then complete the table	the fields of a row add up to the row total, the totals add up to the grand total (1 or 100 %)
“State the businesslike meaning of the quantities ...”	say in words what a set means for the company	e.g. $\bar{M} \cap R$: “a customer who is not a member and redeems the coupon”
expected count	the count of a field if the two properties were independent	$E = \frac{\text{row total} \cdot \text{column total}}{\text{grand total}}$, compare it with the observed count
binomial distribution	the number of successes in n trials	$P(X = k) = \binom{n}{k} p^k (1 - p)^{n-k}$. Requirements: fixed n , two outcomes per trial, the same p in every trial, independent trials
“exactly / not more than / at least ... but not more than ...”	which values of X an event contains	$P(X = k)$, $P(X \leq k)$, $P(a \leq X \leq b)$ with both endpoints. “Fewer than” and “more than” leave out the endpoint
expected value	the long-run average of a random variable	$E[X] = \sum x_i \cdot P(X = x_i)$. Binomial: $\mu = n \cdot p$ and $\sigma = \sqrt{n \cdot p \cdot (1 - p)}$
“with a likelihood of at least ...”	the probability must reach the given value	$P(\dots) \geq$ the value, e.g. $1 - (1 - p)^n \geq 0.95$ for “at least one”
minimum number (of trials)	the smallest whole number n that meets a condition	solve with logarithms, flip the sign when dividing by $\ln(1 - p) < 0$, round up and check $n - 1$ and n

In set-builder notation, $\{x \in \mathbb{R} \mid \dots\}$ and $\{x \in \mathbb{R} : \dots\}$ mean the same.

Set B: Practice Mock (180 minutes, 100 points)

Problem 1: Wrong Orders at a Drive-Through [30 pts. total]

A fast-food restaurant takes 200 orders at its drive-through on every lunch shift. Customers either order in advance with the restaurant’s app (A) or at the speaker in the

drive-through lane ($\bar{A}$). An order is wrong (W) if an item is missing or is not what the customer ordered. On 14 randomly chosen lunch shifts, the numbers of wrong orders were

18, 13, 21, 16, 11, 19, 15, 24, 14, 17, 20, 12, 16, 8

25 % of all orders are app orders, and 2 % of the app orders are wrong.

a) **Determine** the median, the quartiles Q_1 and Q_3 and the interquartile range of the number of wrong orders per shift. On a 15th shift, a new employee worked at the speaker, and 29 orders were wrong. **Decide** with the fences whether this shift is an outlier compared with the 14 shifts. [5 pts.]

b) **Compute** the mean and the sample standard deviation of the number of wrong orders per shift. **Use** the mean to estimate the probability that an order is wrong.

(for verification: $P(W) = 8\%$) [4 pts.]

c) The shift manager claims that one in ten lane orders is wrong. **Draw** a tree diagram with the type of order in the first stage and the unknown $x = P(W | \bar{A})$ on the branch from $\bar{A}$ to W . **Prove** the manager's claim with the law of total probability. [4 pts.]

d) An assistant says: "App orders are a quarter of all orders, so they are also a quarter of the wrong orders." **Compute** the probability that a wrong order is an app order and **assess** the assistant's statement. **Check** with $P(A \cap W) = P(A) \cdot P(W)$ whether the events A and W are independent, and **explain** what the result means for the restaurant. [5 pts.]

e) On an evening shift, the drive-through takes 150 orders. In the evening, 60 % of the orders are app orders. As at lunch, 2 % of the app orders are wrong, but 15 % of the lane orders are wrong. The restaurant puts an extra employee at the pickup window on every shift on which more than 12 wrong orders are expected. **Compute** the probability that an order on an evening shift is wrong and the expected number of wrong orders per evening shift. **Decide** for the lunch shift and for the evening shift whether the extra employee is needed. [3 pts.]

f) On one lunch shift, a trainer checks 22 lane orders. Each lane order is wrong with probability 10 %, independently of the other orders. **Compute** the probability that

(1) exactly 2 of the checked orders are wrong,

(2) fewer than 2 of the checked orders are wrong,

(3) more than 3 of the checked orders are wrong.

[5 pts.]

g) A mystery shopper (a person who is paid to test the service like a normal customer) orders with the app, on different days. **Determine** how many app orders the mystery shopper must place to receive at least one wrong order with a likelihood of at least 90 %. [4 pts.]

Problem 2: Testing Honey at an Importer [35 pts. total]

An importer buys honey in barrels from beekeepers' cooperatives abroad and bottles it for supermarket chains. Some barrels are mixed with cheap sugar syrup (M). The importer tests every barrel with a quick test. Barrels with a positive result (+) go back to the supplier, barrels with a negative result (−) are treated as pure honey and are delivered to the supermarkets. The prevalence (base rate) of barrels mixed with syrup is 3 %, the test's sensitivity (true positive rate) is 94 % and its specificity (true negative rate) is 96 %.

- a) **Create** a tree diagram for the testing and **complete** the following table for 20,000 barrels. [7 pts.]

	positive (+)	negative (−)	sum
mixed with syrup (M)			
pure honey ($\bar{M}$)			
sum			20,000

(for verification: 1,340 of the 20,000 barrels test positive)

- b) **Compute** the probability that a barrel with a positive test result does indeed contain syrup. **Explain** why this probability is so much lower than the sensitivity of 94 %. [4 pts.]
- c) **Compute** the probability that a barrel contains syrup but is incorrectly identified as pure honey. **Compute** the negative predictive value, i.e. the share of the delivered barrels that really are pure honey. [4 pts.]
- d) Consider a random sample of 140 barrels from all barrels that arrive, before the test. **Compute** the probability for the following events:

- (1) Exactly 3 barrels are mixed with syrup.
- (2) Not more than 6 barrels are mixed with syrup.
- (3) At least 4 but not more than 10 barrels are mixed with syrup.

[5 pts.]

- e) Before the quick test existed, the importer sent barrels to a laboratory one by one. **Determine** the number of barrels which need to be tested to find at least one barrel mixed with syrup with a likelihood of 97 %. [4 pts.]
- f) The importer works with 10 cooperatives. This season, 4 of them delivered at least one barrel mixed with syrup. A food inspector visits 3 of the 10 cooperatives, picked at random. **Give** the number of ways to pick the 3 cooperatives. **Compute** the probability that the inspector's visits include exactly 2 of the cooperatives that delivered syrup, and the probability that the inspector visits at least one cooperative that delivered syrup. [5 pts.]
- g) For their own brand, the supermarket chains accept the honey only if at most 1 of 1,000 delivered barrels contains syrup. Only the sensitivity of the test can be

changed. **Find** the smallest sensitivity the test needs and the largest false negative rate that goes with it.

Hint: The denominator of a probability is positive, so multiplying an inequality by it keeps the inequality sign.

[6 pts.]

Problem 3: Season Tickets and Interval Drinks at a Concert Hall [35 pts. total]

A concert hall sells season tickets (one ticket for all concerts of a season) and single tickets. 2,000 visitors attend one concert. Let T be the event “the visitor has a season ticket” and B the event “the visitor buys a drink at the bar during the interval”. 30 % of the visitors buy a drink. Of the visitors who buy a drink, 60 % have a season ticket. 42 % of all visitors have no season ticket and buy no drink.

- a) **Display** the fractions in a contingency table. **Fill in** the blank fields. Then **give** the table of counts for the 2,000 visitors. [7 pts.]

	season ticket (T)	single ticket ($\bar{T}$)	sum
buys a drink (B)			
buys no drink ($\bar{B}$)			
sum			100 %

(for verification: 46 % of the visitors have a season ticket)

- b) **Compute** with Bayes’ theorem the probability that a visitor with a season ticket buys a drink. **Compute** the same probability for a visitor with a single ticket. [5 pts.]
- c) The bar manager thinks that having a season ticket and buying a drink have nothing to do with each other. **Check** this with the expected count of the field $T \cap B$. **Explain** what the result means for planning the bar. [4 pts.]
- d) **Argue** whether the following statements are true or false.

(1) If the bar manager were right, 324 visitors with a single ticket would buy a drink.

(2) 58 % of the visitors have a season ticket or buy a drink.

(3) 60 % of the visitors with a season ticket buy a drink.

(4) The events “has a season ticket” and “buys no drink” are mutually exclusive.

[4 pts.]

- e) For a survey about new drinks, the bar manager wants to interview visitors who have a single ticket and buy a drink. At the next concert, the manager speaks to randomly chosen visitors one after another. **Set up** an inequality for the number n of visitors the manager has to speak to so that, with a likelihood of at least 95 %, at least one of them has a single ticket and buys a drink, and **solve** it. **Explain** why rounding to the nearest whole number would give a wrong answer here. [4 pts.]

At the next concert, 20 visitors are chosen at random. X is the number of them who buy a drink during the interval. The shares from a) stay the same.

- f) **State** the four requirements of a binomial distribution and **explain** for each of them whether it holds for X . **Compute** the expected value and the standard deviation of X , and **interpret** the expected value. [5 pts.]

- g) **Compute** the probability for the following events:

- (1) Exactly 5 of the 20 visitors buy a drink.
 (2) Not more than 8 of the 20 visitors buy a drink.
 (3) At least 5 but not more than 11 of the 20 visitors buy a drink.

[6 pts.]

Exam Preparation Checklist

Before the exam, check that you can:

- ☐ Compute the mean, the median, the quartiles, the interquartile range and the sample standard deviation (with $n - 1$) from raw data, and find outliers with the fences (1.5 IQR below Q_1 and above Q_3)
- ☐ Read a relative frequency or a mean per day as a probability, with all items as the base
- ☐ Use the 68-95-99.7 rule, and z-scores with the values of the standard normal distribution that the task gives
- ☐ Use the complement rule (“at least one” = $1 - \text{“none”}$), the addition rule and the test $P(A \cap B) = P(A) \cdot P(B)$ for independent events, and tell mutually exclusive events from independent ones
- ☐ Count selections with the counting principle, permutations and combinations, and compute “exactly k ” and “at least one” by counting
- ☐ Draw a tree diagram with branch and path probabilities, and check that the branches at every point and all path probabilities add up to 1
- ☐ Use the law of total probability, also with an unknown rate as a linear equation (“show that ...”)
- ☐ Find the base of every conditional probability: the group after “of the ...”, “among the ...” or “given that ...” goes behind the bar
- ☐ Compute the positive and the negative predictive value with Bayes’ theorem, and tell a joint probability (“... but is incorrectly identified as ...”) from a conditional one
- ☐ Determine a maximum error rate or a minimum base rate for a required predictive value with an inequality
- ☐ Complete a contingency table from percentages (“Fill in the blank fields”) and check that every row and every column adds up to its total and the totals to the grand total
- ☐ Test independence with $P(A \cap B) = P(A) \cdot P(B)$, with $P(A | B) = P(A)$ and with the expected count, and interpret the result
- ☐ Argue whether statements are true or false, with complements such as “neither” = $\overline{A \cup B}$ and “not both” = $\overline{A \cap B}$

- ☐ Check the four requirements of the binomial distribution and compute $\mu = n \cdot p$ and $\sigma = \sqrt{n \cdot p \cdot (1 - p)}$
- ☐ Translate “exactly”, “not more than”, “at least ... but not more than ...”, “fewer than” and “more than” into values of X with the correct endpoints, and use the complement when it is shorter
- ☐ Get binomial values with the calculator: $P(X = x)$ with Binomial-Dichte and $P(X \leq x)$ with the cumulative distribution (**MENU** → **7** → **▼** → **1** → **2**), “at least” as $1 - P(X \leq k - 1)$ and “between” as $P(X \leq b) - P(X \leq a - 1)$, with the decimal comma (**0,03**)
- ☐ Tell the binomial from the hypergeometric distribution (a draw without replacement from a small lot), and use the geometric distribution for the first success
- ☐ Determine a minimum number n with logarithms: flip the sign when dividing by $\ln(1 - p) < 0$, round up and check $n - 1$ and n

 Tip

Show all work, name the base of every probability, and interpret your results in the business context.