1. Eine Einkommensstatistik für Vierzigjährige (Jahresnettoeinkommen, in 1000 Euro, klassiert), zeigte folgende Anzahlen (absolut) für die Einkommen in den jeweiligen Klassen, getrennt für Personen ohne und solche mit Hochschulabschluss (HA):

	10 bis 20	**20** – 30	**30** – 40	**40** bis 75
ohne HA	40	65	30	25
mit HA	15	25	30	70

a. Stellen Sie die **Einkommensverteilung der *Personen ohne HA*** durch eine **geeignete Verteilungsfunktion** dar!

b. Stellen Sie die **Einkommensverteilung der *Personen mit HA*** durch ein **Histogramm** dar!

c. Berechnen Sie – soweit wie möglich – das ***mittlere Einkommen aller Personen***.

d. Geben Sie das **Medianeinkommen** der ***Personen mit HA*** an!

4 + 4 + 1 + 1

a. Approximierende VF: Knickpunkte (10, 0) (20, 0.25) (30, 0.656) (40, 0.844) und (75, 1)

b. Histogramm: Höhen über den Intervallen: 1.5, 2.5, 3, 2

c. 35.46 d. 40

2. Im Rahmen eines Bewerbungsverfahrens mussten 8 zufällig ausgewählte Kandidaten Zwölf Schlussrechnungen lösen. Die Anzahl der richtigen Lösungen wurde notiert.

In der Zeile darunter die Antworten dieser Personen auf die Frage „Wie hoch schätzen Sie Ihre Fähigkeit zum analytischen Denken ein", klassifiziert nach G(ering), M(ittel), H(och), (S)uper:

Kandidat Nr.	1	2	3	4	5	6	7	8
Richtige Lösungen:	12	6	7	10	6	11	8	6
Selbsteinschätzung	M	H	H	S	M	H	M	G

a. **Berechnen** Sie eine geeignete Kennzahl um den Zusammenhang zwischen der Anzahl richtig gelöster Aufgaben und Selbsteinschätzung zu beschreiben und **interpretieren** Sie die erhaltene Zahl!

b. Wie könnte man die folgende Frage beantworten:
„Lässt sich auf Grund der vorliegenden Daten zeigen, dass bei Kandidaten mit höherer Selbsteinschätzung auch die Leistung besser ist?"

- ***Welcher Test*** ist dazu geeignet? ***Einseitig*** oder ***zweiseitig***?
- Formulieren Sie ***Null- und Gegenhypothese in Worten***
- Formulieren Sie Null- und Gegenhypothese auch in Formeln!

Dieser Test ist ***nicht*** ***durchzuführen***!!!

6 + 4

a. Rangkorrelationskoeffizient r_s (ordinalskaliertes Merkmal 2)

Summe der di² = 56 (bei gegengleicher Rangzuweisung: 115)

Spearman rs = 0.3333 (**rs** = – 0.369)

Positiver Zusammenhang, nicht sehr ausgeprägt:

b. Test auf positive Korrelation im Sinne von Rangkorrelation, einseitig,

H0: Selbsteinschätzung und Leistung sind nicht (oder sogar negativ) korreliert

H1: Zwischen Selbsteinschätzung und Leistung liegt positive Korrelation vor

H0: „korr" <= 0 H1: korr >0

3. a. Bestimmen Sie für eine ***poissonverteilte*** Zufallsgröße ***X*** mit dem Erwartungswert 1.2 die Wahrscheinlichkeit $P(X \geq 3)$

b. Eine andere Zufallsgröße ***Y*** sei ***N(7, 2.5)***-verteilt.

b1 Zeichnen Sie die zugehörige Dichtefunktion (Skala auf der x-Achse!)

b2 Zeichnen Sie die zugehörige Verteilungsfunktion!

b3 Bestimmen Sie $P(4 \leq Y \leq 9.5)$

b4 Für welche Zahl c gilt: $P(Y < c) = 0.18$

c. Veranschaulichen Sie die Aufgabe b4 in Ihren beiden Zeichnungen!

2 + 6 + 2

a. $\mathbf{P(X \geq 3)} = 1 - P(\{0, 1, 2\}) = \mathbf{0.1205}$

b. b1: Glockenkurve

b2: Flache S-förmige Kurve mit Werten zwischen 0 und 1

b3: $\mathbf{P(4 \leq Y \leq 9.5) = 0.7262}$

b4: $\Phi((c-7)/2.5) = -0.9154$ **c = 4.7115**

c. in b1 aus der **Fläche**, in b2 aus der **Höhe**

4. Unter 7160 befragten Männern waren Raucher/Nichtraucher wie angegeben nach Altersgruppen verteilt:

	Raucher	Nichtraucher
Unter 40	1865	1888
Über 40	1778	1629

a. Bestimmen Sie ein **zweiseitiges Konfidenzitervall** für den Anteil der Raucher unter den Jüngeren (unter 40-Jährigen) Männern, und zwar zum Konfidenzniveau 99 Prozent!
b. Lässt sich aus den Angaben auf unterschiedliches Rauchverhalten von jüngeren und älteren Männern schließen?
Führen Sie einen **geeigneten Test** zu den zwei Testniveaus 0.1 und 0.01 durch und **beantworten Sie die gestellte Frage** so genau wie möglich!

4 + 6

a. Schätzwert für p: 0.4969 = Anteil der Raucher, nur bei den 3753 Jüngeren

a = 0.021
Konf$\{0.4759 \le p \le 0.5179\}$

b. Vierfeldertest:

1865	1888	**3753**
1778	1629	**3407**
3643	**3517**	***7160***

Testgröße = 4.44

Niveau 0.1: Tabellenwert 2.707 H0 ablehnen, **Abhängigkeit *signifikant***

Niveau 0.01: Tabellenwert 6.63 H0 beibehalten
Abhängigkeit *nicht hochsignifikant*

5. Unter 8 geprüften Bauteilen aus einer umfangreichen Produktion war genau einer defekt. Lässt sich damit die Behauptung
"Der Ausschussanteil der Gesamtheit liegt unter dreißig Prozent" signifikant bestätigen?
(Niveau α **= 0.1**)

a. **Bestimmen sie dazu den *kritischen Bereich*** für die Anzahl der defekten Stücke unter den 8 geprüften.

b. **Wie ist demnach zu entscheiden**?

c. Wie groß ist **die WK für einen Fehler 2. Art**, wenn der tatsächliche Ausschussanteil ***vierzig*** Prozent beträgt?

4 + 3 + 3

Binomialtest

a. H1: $p < 0.3$ H0: $p \geq 0.3$

$K = \{0, 1, \ldots, c\}$ mit $F(c) < 0.1 \leq F(c+1)$

b. $\mathbf{K = \{0\}}$ $\mathbf{t0 = 1}$ H0 beibehalten, Behauptung **nicht „bestätigt"**

c. Ist $p_{wahr} = 0.4$, dann ist die Nullhypothese wahr, ein Fehler 2. Art kann nicht gemacht werden. P(Fehler 2. Art) = 0

6. Letzte Woche hatten in Graz elf männliche Neugeborene ein mittleres Gewicht von 3357 g, bei einer Stichprobenstandardabweichung von 458g. Bei 17 neugeborenen Mädchen waren die Zahlen 3040 und 345.
Die Gewichte werden als normalverteilt betrachtet.

a. **Kann man „Gleichheit der Varianzen" akzeptieren?**
(Testen Sie zum Niveau $\alpha = 0.05$)

b. **Testen** sie zum Niveau 0.05 die Gegenhypothese
H_1 : Neugeborene Mädchen sind leichter als Buben
Und interpretieren Sie das Ergebnis genau!

c. Wie groß darf die Differenz der zwei Stichprobenmittelwerte höchstens sein, damit die Nullhypothese zu b. beibehalten werden kann?

3 + 5 + 2

a. F-Test Fisher x…Buben, y….. Mädchen
H0: gleiche Varianzen

Testgröße t0 = **1.76**

Tabellenwert zu $1-\alpha/2 = 0.975$: ***$F_{10, 16} = 3.0$*** H0 beibehalten

b. Zweiproben-t-Test, einseitig (y…Buben, x…. Mädchen)
H1: $\mu_y > \mu_x$ s = 392.3

Testgröße **t0 = 2.0881** ***Tabellenwert t(0.95, 26) = 1.706***
H1 ist signifikant. „Buben sind schwerer"

c. für welche Differenz wird H0 noch beibehalten: Bis zu **d krit = 259**