Klausur: **Statistik**

Jürgen Meisel

Zugelassene Hilfsmittel: nicht progr. Taschenrechner
Bearbeitungszeit: **60 Minuten**

Anmerkung zur Bearbeitung:
Die Klausur besteht aus insgesamt 6 Aufgaben. Sie müssen nur 5 davon bearbeiten.
Bitte wählen Sie 5 Aufgaben aus und streichen Sie die nicht bearbeitete Aufgabe bitte auf dem Aufgabenblatt.

Aufgabe 1: Mittelwerte und Streumaße

Aufgabe 2: Preisindizes

Aufgabe 3: Lineare Regression und Korrelation

Aufgabe 4: Konzentrationsprozesse

Aufgabe 5: Baumdiagramm, Pfadregeln und Binomialverteilung

Aufgabe 6: Normalverteilung

Teilbereich Statistik

1.) Mittelwerte und Streumaße: Ergebnisse Abiturprüfung 2012 | **20** | |

Gegeben sei eine Ergebnisübersicht der Abiturprüfung Mathematik von 2012:

Schülernummer	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
Ergebnis Abitur	12	05	12	07	06	07	09	08	03	07	02	03	03	13	09

Anmerkung:
Die Ergebnisse sind im 15-Punkte-System der gymnasialen Oberstufe gegeben.

a) Ermitteln Sie das arithmetische Mittel.

Lösung:

Ergebnis	02	03	04	05	06	07	08	09	10	11	12	13
Anzahl	1	3	0	1	1	3	1	2	0	0	2	1

Mittelwert:

$$\mu = \frac{1\cdot 2+3\cdot 3+1\cdot 5+1\cdot 6+3\cdot 7+1\cdot 8+2\cdot 9+2\cdot 12+1\cdot 13}{15} = \frac{106}{15} = 7{,}067$$

b) Berechnen Sie die Varianz und die Standardabweichung.

Lösung:

Varianz: $\sigma^2 = \frac{922}{15} - 7{,}067^2 = 11{,}5289 \rightarrow \sigma = 3{,}395$

c) Bestimmen Sie nun noch den Median.

Lösung: Median: $\overline{x_M} = x_{\frac{n+1}{2}} \rightarrow \overline{x_M} = x_8 = 7$

d) Erstellen/zeichnen Sie nun einen kompletten Boxplot mit einer Berechnung der Quartile 1 und 3.

Lösung:

Quartile 1: $\overline{x_{0,25}} = x_{[n \cdot p]+1} \rightarrow \overline{x_{0,25}} = x_{[15 \cdot 0,25]+1} = x_4 = 3$

Quartile 3: $\overline{x_{0,75}} = x_{[n \cdot p]+1} \rightarrow \overline{x_{0,75}} = x_{[15 \cdot 0,75]+1} = x_{12} = 9$

=> Zeichnung Boxplot

2.) Preisindizes

20	

Familie Winzig leistet sich leistet sich folgende Produkte und Ausgaben. Für die einzelnen Jahre wurde aus dem Haushaltsbuch folgende Übersicht erstellt:

	Preis 2008	Umsatz 2008	Preis 2011	Verbrauch 2011
Chips	1,00 € / Tüte	40,00 €	1,50 € / Tüte	50 Tüten
Erdnüsse	6,00 € / kg	36,00 €	9,00 € / kg	7 kg
Bier	4,20 € / ltr.	504,00 €	5,50 € / ltr.	150 ltr.

a) Berechnen Sie die Preisindizes nach Laspeyres und Paasche.

Lösung:

$$L_P = \frac{\sum p_{1i} \cdot q_{0i}}{\sum p_{0i} \cdot q_{0i}} \rightarrow \frac{1{,}50 \cdot 40 + 9{,}00 \cdot 6 + 5{,}50 \cdot 120}{40 + 36 + 504} = \frac{774}{580} = 1{,}3345$$

$$P_P = \frac{\sum p_{1i} \cdot q_{1i}}{\sum p_{0i} \cdot q_{1i}} \rightarrow \frac{1{,}5 \cdot 50 + 9{,}00 \cdot 7 + 5{,}5 \cdot 150}{1{,}00 \cdot 50 + 6{,}00 \cdot 7 + 4{,}20 \cdot 150} = \frac{963}{722} = 1{,}3338$$

b) Wie groß ist der durchschnittliche jährliche Preisanstieg in %, wenn der Preisindex nach Laspeyres zugrunde gelegt wird?

Lösung: $g = \sqrt[3]{1{,}3345} = 1{,}10095 \rightarrow 10{,}1[\%]$

3.) Lineare Regression und Korrelation

20	

Gegeben sei eine Ergebnisübersicht der Abiturprüfung Mathematik von 2012. Zudem sind noch die Vornoten gegeben.

Im Raum steht die Behauptung, dass die Vornoten und die Ergebnisse der Abiturprüfung relativ stark voneinander abweichen.

Schülernummer	**Vornote (x)**	**Ergebnis Abitur (y)**
1	15	12
2	04	05
3	15	12
4	10	07
5	14	13
6	09	09
7	10	08
8	04	02
9	07	03
10	07	07

Anmerkung:
Die Ergebnisse sind im 15-Punkte-System der gymnasialen Oberstufe gegeben.

Ermitteln Sie die Regressionsgerade $y = b_0 + b_1 x$ und den Korrelationskoeffizienten und nehmen Sie aufgrund Ihrer Ergebnisse **kritisch** Stellung zu obiger Aussage.

Lösung: Ansatz: $y = b_0 + b_1 x$

$$\mu_x = \frac{1}{n}\sum_{i=1}^{n} x_i = 9{,}5 \qquad \mu_y = \frac{1}{n}\sum_{i=1}^{n} y_i = 7{,}8$$

$$b_1 = \frac{\frac{1}{n}\sum_{i=1}^{n}(x_i y_i) - \mu_x \cdot \mu_y}{\frac{1}{n}\sum_{i=1}^{n}\left(x_i^2\right) - \mu_x^2} \rightarrow \frac{\frac{1}{10}\cdot 871 - 9{,}5\cdot 7{,}8}{\frac{1}{10}\cdot 1.057 - 9{,}5^2} = \frac{13}{15{,}45} = 0{,}8414$$

$$b_0 = \mu_y - b_1 \cdot \mu_x \rightarrow 7{,}8 - 0{,}8414 \cdot 9{,}5 = -0{,}19$$

Regressionsgerade: $y = -0{,}19 + 0{,}8414x$

Korrelationskoeffizient:

20	

$$r = \frac{\sum_{i=1}^{n}(x_i y_i) - n \cdot \mu_x \cdot \mu_y}{\sqrt{\sum_{i=1}^{n}(x_i^2) - n \cdot \mu_x^2} \cdot \sqrt{\sum_{i=1}^{n}(y_i^2) - n \cdot \mu_y^2}} \rightarrow 0{,}9187$$

Gleichartige Entwicklung; aber dies bedeutet nicht, dass die Schüler im Abitur die gleichen Noten erhalten haben wie bei der Vornote; allerdings können natürlich die Noten im Abitur von den Vornoten abweichen – aber die Entwicklung ist bei den Schülern identisch; im Vorfeld gute Schüler haben auch entsprechende Leistungen im Abitur – entgegengesetzte Entwicklungen sind nicht zu erkennen.

4.) Konzentrationsprozesse

Die Wichtel GmbH aus Mannheim ist ein mittelständiges Unternehmen mit 10 Kunden.

Das Unternehmen arbeitet in der Glas- und Kunststoffherstellung.
Für das 1. Quartal 2012 ergaben sich folgende Umsätze:

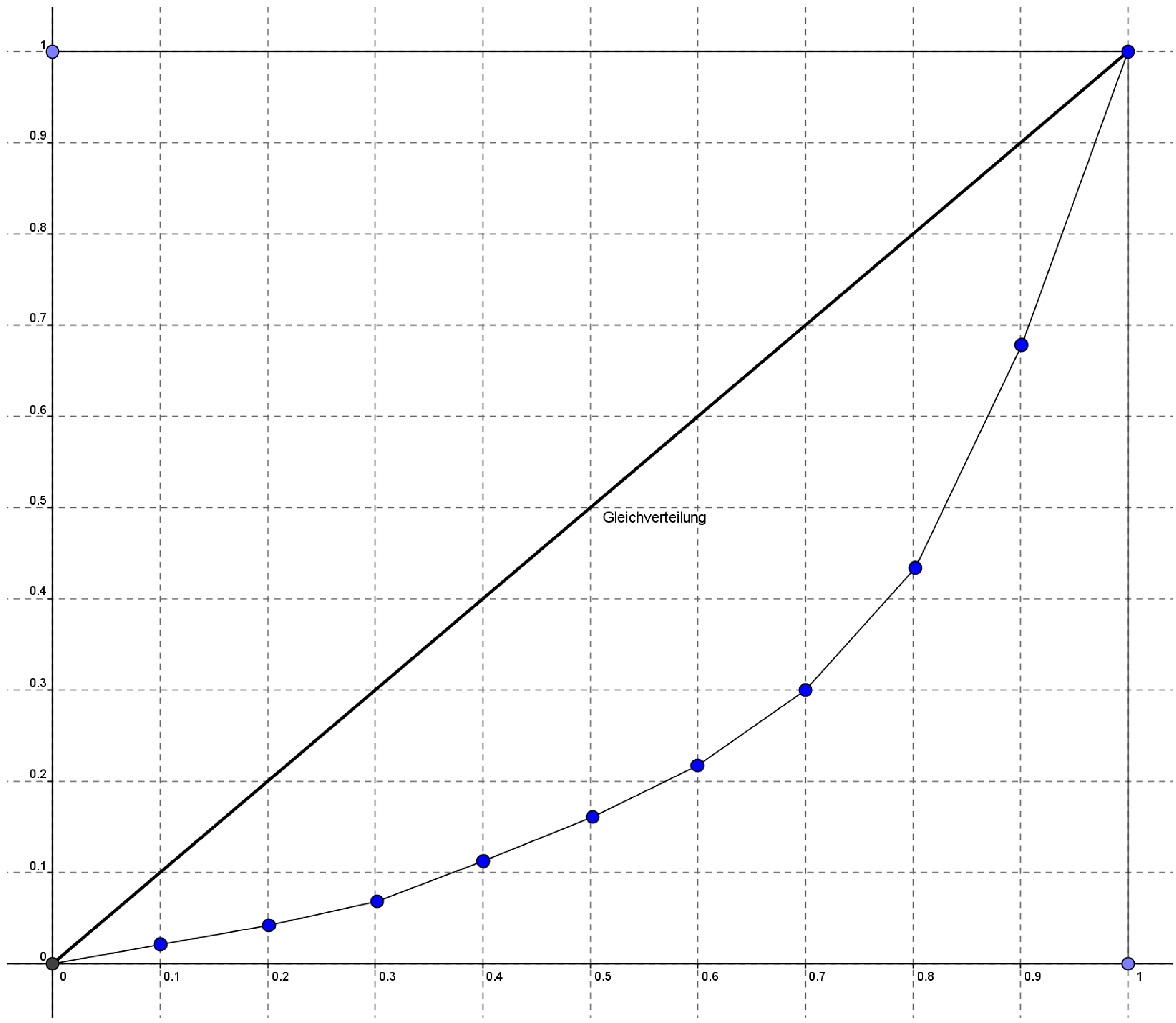
Unternehmen (alphabetisch)	Umsatz
Abels	60.000,00
Eigenmeier	160.000,00
Fisher	65.000,00
Heilbutt	210.000,00
Joker-Wien	100.000,00
Keller-Assel	50.000,00
Küsterberg	315.000,00
Meichelborn	620.000,00
Schneiderle	800.000,00
Zweistein	120.000,00

a) Erstellen Sie die zugehörige Lorenzkurve.

Lösung:

Unternehmen (geordnet nach Umsatz)	Kumuliert Unternehmen	Umsatz	Kumuliert Umsatz
Keller-Assel	1	50.000,00	50.000,00
Abels	2	60.000,00	110.000,00
Fisher	3	65.000,00	175.000,00
Joker-Wien	4	100.000,00	275.000,00
Zweistein	5	120.000,00	395.000,00
Eigenmeier	6	160.000,00	555.000,00
Heilbutt	7	210.000,00	765.000,00
Küsterberg	8	315.000,00	1.080.000,00
Meichelborn	9	620.000,00	1.700.000,00
Schneiderle	10	800.000,00	2.500.000,00

Unternehmen (geordnet nach Umsatz)	Relativ kumuliert	Relativ kumuliert	Flächen
Keller-Assel	0,1	0,0200	0,00100
Abels	0,2	0,0440	0,00320
Fisher	0,3	0,0700	0,00570
Joker-Wien	0,4	0,1100	0,00900
Zweistein	0,5	0,1580	0,01340
Eigenmeier	0,6	0,2220	0,01900
Heilbutt	0,7	0,3060	0,02640
Küsterberg	0,8	0,4320	0,03690
Meichelborn	0,9	0,6800	0,05560
Schneiderle	1,0	1,0000	0,08400



b) Berechnen Sie den Gini-Koeffizient.

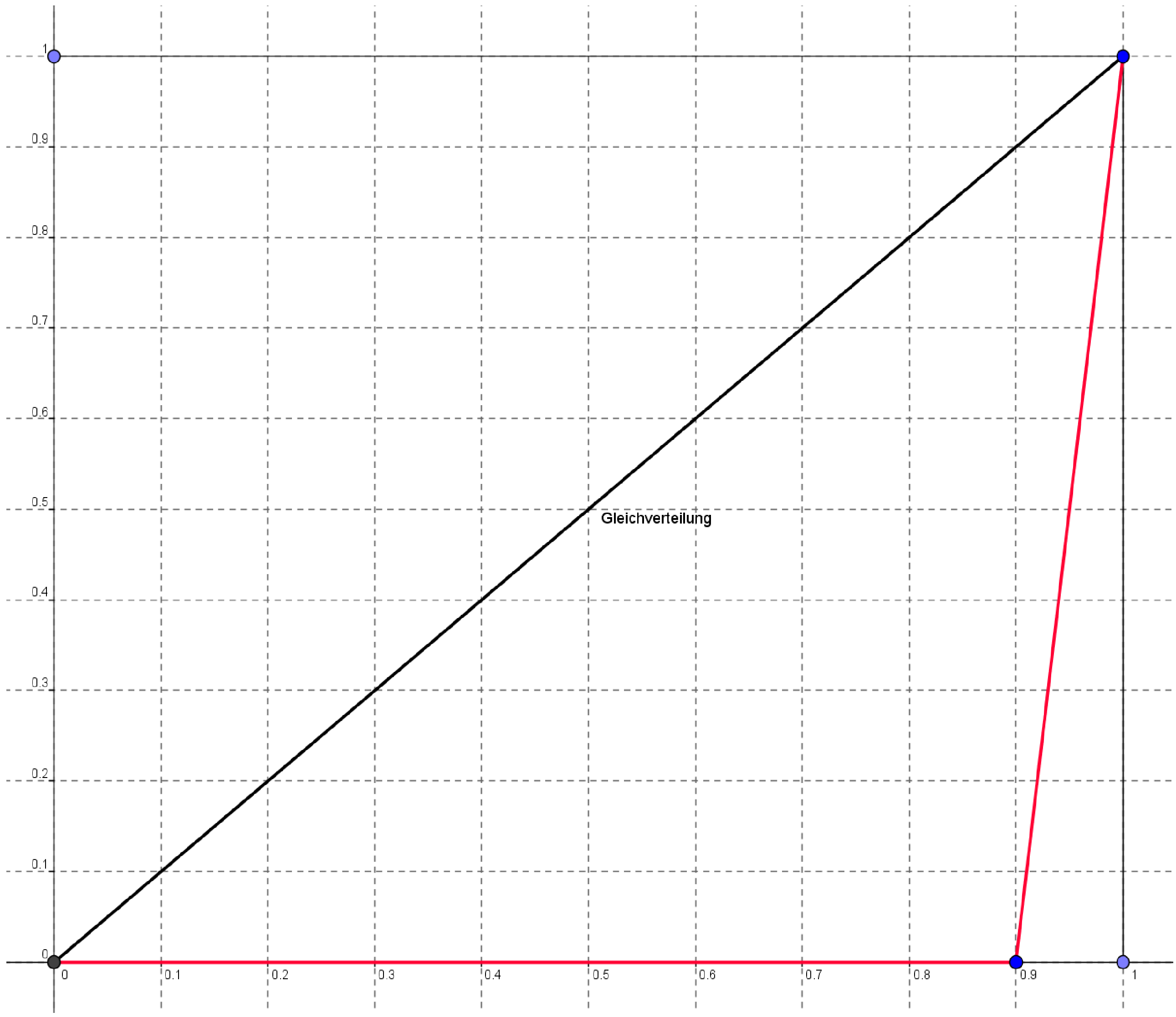
Lösung: *Berechung der Flächen mit relativen kumulierten Häufigkeiten:*

$$Gini = \frac{\text{Konzentrationsfläche K}}{\text{maximale Konzentrationsfläche } K_{max}}$$

Summe (unth. Lorenz):	**0,25420**
K(max):	**0,50**
K:	**0,24580**
Gini:	**0,491600**
Gini (normiert):	**0,546222**

c) Wie würden sich Lorenzkurve und Gini-Koeffizient ändern, wenn die Wichtel GmbH nur noch an die Unternehmung Schneiderle liefern würde und die anderen Unternehmen nicht mehr bedienen könnte?

Lösung:



$$Gini = \frac{0{,}5 - \frac{1}{2} \cdot 0{,}1 \cdot 1}{0{,}5} = \frac{0{,}45}{0{,}5} = 0{,}9$$

$$Gini[normiert] = 1 \rightarrow \text{totale/vollständige Konzentration}$$

Teilbereich Stochastik

5.) **Baumdiagramm, Pfadregeln und Binomialverteilung**

20	

Teil 1:

Ein Prüfling muss drei Klausuren schreiben, von denen er mindestens zwei bestehen muss. Besteht er alle drei, so besteht er „mit Auszeichnung".
Teil A besteht er mit 90 %, Teil B mit 95 %. Bei Teil C – sein Problemfach – fällt er mit einer Wahrscheinlichkeit von 35 % durch.
Wie groß ist die Wahrscheinlichkeit, dass er die Prüfung

a) „mit Auszeichnung" besteht

Lösung:

$$P(\text{"}mit\ Auszeichung\text{"}) = 0{,}9\cdot 0{,}95\cdot 0{,}65 = 0{,}55575 = 55{,}575\,[\%]$$

b) besteht, aber ohne Auszeichnung?

Lösung:

$$P(\text{"}bestehen\text{"}) = 0{,}1\cdot 0{,}95\cdot 0{,}65+0{,}9\cdot 0{,}05\cdot 0{,}65+0{,}9\cdot 0{,}95\cdot 0{,}35$$

$$P(\text{"}bestehen\text{"}) = 0{,}06175+0{,}02925+0{,}29925 = 0{,}39025 = 39{,}025\,[\%]$$

Teil 2:

Der Inhaber eines kleinen Reisebüros, Rudi Rastlos, weiß aus langjähriger Erfahrung, dass 40 % seiner Kunden den Süden mit Sonne und Meer als Reiseziel bevorzugen.

Mit welcher Wahrscheinlichkeit befinden sich unter den nächsten 100 Buchungen

(i) mindestens 40 für den Süden?
(ii) höchstens 55 für den Süden?
(iii) zwischen 35 und 50 für den Süden?
(iv) mehr als 60 für den Süden?

Lösung:

(i) $B(X\geq 40) = 1-B(X\leq 39) = 1-0{,}4621 = 0{,}5379$

(ii) $B(X\leq 55) = 0{,}9991$

(iii) $B(35\leq X\leq 50)=B(X\leq 50)-B(X\leq 34)=0{,}9832-0{,}1303 = 0{,}8529$

(iv) $B(X>60) = 1-B(X\leq 60) = 1-1 = 0$

6.) Normalverteilung

20	

Eine Fluggesellschaft bietet Linienflüge mit einem Airbus (**250 Sitzplätze**) an. Erfahrungsgemäß erscheinen nur 90% der Passagiere, die einen Platz gebucht haben, auch tatsächlich zum Abflug.

a) Erklären Sie, warum der Erwartungswert $\mu = 225$ und die Standardabweichung $\sigma = \sqrt{22,5}$ betragen?

Lösung:

Erwartungswert: $\mu = n \cdot p \rightarrow \mu = 250 \cdot 0,9 = 225$

Varianz:

$$\sigma^2 = n \cdot p \cdot (1-p) \rightarrow \sigma^2 = 225 \cdot 0,1 = 22,5$$

$$\rightarrow \sigma = \sqrt{225 \cdot 0,1} = \sqrt{22,5}$$

b) Wie groß ist die Wahrscheinlichkeit, dass bei einem ausgebuchten Flug mindestens 230 Plätze belegt werden?

Lösung:

$$P(X \geq 230) = 1 - P(X < 230) = 1 - \Phi(1,05) = 1 - 0,853 = 0,147$$

$$Umrechnung: \quad x = \frac{230-225}{\sqrt{22,5}} = 1,054$$

Aus ökonomischen Gründen ist die Fluggesellschaft dazu übergegangen, die Flüge überbuchen zu lassen.

c) Wie groß ist die Zahl der gebuchten Plätze, wenn man eine 10%ige Überbuchung zugelassen hat?

Lösung: 275 Plätze

d) Wie groß ist die Wahrscheinlichkeit, dass bei einer 10%igen Überbuchung nicht alle erscheinenden Fluggäste transportiert werden können (d.h. dass mindestens 251 Passagiere kommen)?

Lösung:

$$P(X \geq 251) \quad = \quad 1 - P(X \leq 250) \quad = \quad 1 - \Phi(0{,}50) \quad = \quad 1 - 0{,}691 \quad = \quad 0{,}309$$

$$Umrechnung: \quad x \quad = \quad \frac{250 - 247{,}5}{\sqrt{24{,}75}} \quad = \quad 0{,}5025$$

$$neuer\ Erwartungswert\ und\ neue\ Varianz:$$

$$\mu \quad = \quad 275 \cdot 0{,}9 \quad = \quad 247{,}5$$

$$\sigma^2 \quad = \quad 247{,}5 \cdot 0{,}1 \quad = \quad 24{,}75 \quad \rightarrow \quad \sigma \quad = \quad \sqrt{24{,}75}$$