Klausur Statistik

Name:

Matrikel-Nr.:

Studiengang: ☐ WI ☐ AI

Prüfer:	Prof. Dr. Martin Hulin
Dauer:	60 Minuten
Datum:	3. Juli 2008
Hilfsmittel:	Alle im Prüfungsplan angegebenen Hilfsmittel
Kennzahlen:	3609 und 4110

Dokumentieren Sie Ihren Lösungsweg durch Angabe aller Zwischenschritte, sonst gibt es keine Punkte!

Lösen Sie die Aufgaben bitte direkt auf dem Aufgabenblatt.
Rückseiten sind bedruckt!

Bewertung:

Aufgabe	1 (10)	2 (21)	3 (19)			
Punkte						
Summe	(50)					

Tabelle 1: Inverse der Verteilungsfunktion für die Standardnormalverteilung

x	$\Phi^{-1}(x, \mu = 0, \sigma = 1)$	x	$\Phi^{-1}(x, \mu = 0, \sigma = 1)$
0,025	-1,95996	0,525	0,0627068
0,05	-1,64485	0,55	0,125661
0,075	-1,43953	0,575	0,189118
0,1	-1,28155	0,6	0,253347
0,125	-1,15035	0,625	0,318639
0,15	-1,03643	0,65	0,38532
0,175	-0,934589	0,675	0,453762
0,2	-0,841621	0,7	0,524401
0,225	-0,755415	0,725	0,59776
0,25	-0,67449	0,75	0,67449
0,275	-0,59776	0,775	0,755415
0,3	-0,524401	0,8	0,841621
0,325	-0,453762	0,825	0,934589
0,35	-0,38532	0,85	1,03643
0,375	-0,318639	0,875	1,15035
0,4	-0,253347	0,9	1,28155
0,425	-0,189118	0,925	1,43953
0,45	-0,125661	0,95	1,64485
0,475	-0,0627068	0,975	1,95996
0,5	0		

Aufgaben **(Punkte)**

1. Ein Computer kostet in einem Computer-Laden 500 € Barpreis. Alternativ ist eine Zahlung in zwei gleich hohen Raten möglich, die erste Rate sofort, die zweite nach einem Jahr. Um die Rate ***r*** zu bestimmen, rechnet der Händler mit einer jährlichen Verzinsung von 5%. Außerdem weiß er aus Erfahrung, dass die zweite Rate nur von 90% der Kunden bezahlt wird. Bestimmen Sie die Rate ***r*** so, dass der Erwartungswert für den Wert der Einnahmen nach einem Jahr bei Ratenzahlung gleich dem Wert der Einnahmen nach einem Jahr bei Barzahlung ist. **(10)**

2. Im Unterrichtsfach Geschichte haben 4 Schüler bei einer Klassenarbeit, in der Jahreszahlen abgefragt wurden, folgende Noten erzielt. Zusätzlich ist die Lernzeit der Schüler angegeben.

Schüler-Nr.	Lernzeit [h]	Note
1	1	4,3
2	3	2,0
3	4	1,0
4	2	4,0

a) Zeichnen Sie die Ergebnisse der Stichprobe in ein Streudiagramm ein. Beschriften Sie auch die Achsen **(5)**

b) Wie groß ist der Korrelationskoeffizient? Sie brauchen ihn dazu nicht auszurechnen, sondern können ihn auch schätzen. Kreuzen Sie das richtige Kästchen an! **(3)**

☐ 0 ☐ 100 ☐ -1 ☐ +1 ☐ 1,5 ☐ -1,5 ☐ 95 ☐ -95

☐ 0,97 ☐ -0,97 ☐ 0,24 ☐ -0,24

c) Kreuzen Sie alle Begründungen an, die Ihre Wahl bei b) stützen. Setzen Sie maximal vier Kreuze. **(4)**

☐ Der Korrelationskoeffizient ist die Steigung der Regressionsgeraden

☐ Der Korrelationskoeffizient ist eine Zahl zwischen 0 und 1

☐ Der Korrelationskoeffizient ist eine Zahl zwischen -1 und +1

☐ Der Korrelationskoeffizient ist eine Zahl zwischen 0 und 100

☐ Der Korrelationskoeffizient ist eine Zahl zwischen -100 und +100

☐ Der Korrelationskoeffizient liegt nahe bei +1 oder -1, wenn die beiden Merkmale stark voneinander abhängen.

☐ Der Korrelationskoeffizient ist negativ, wenn die Steigung der Regressionsgeraden negativ ist.

☐ Der Korrelationskoeffizient kann nur dann 100 sein, wenn alle Punkte der Stichprobe exakt auf der Regressionsgeraden liegen.

☐ Der Korrelationskoeffizient kann nur dann 1 sein, wenn alle Punkte der Stichprobe exakt auf der Regressionsgeraden liegen.

☐ Der Korrelationskoeffizient ist 0, wenn alle Punkte der Stichprobe exakt auf der Regressionsgeraden liegen.

☐ Der Korrelationskoeffizient kann beliebig groß werden, wenn die Punkte weit verstreut liegen, da er die Summe der Abweichungs-Quadrate von der Regressionsgeraden ist.

d) Zwei Studenten, Hans und Thomas, berechnen die Regressionsgerade. Sie ermitteln zwei verschiedene Regressionsgeraden.
Hans: $\boxed{Note = 7{,}0 - 1{,}5 \cdot Lernzeit}$ **Thomas:** $\boxed{Note = 5{,}0 - 1{,}0 \cdot Lernzeit}$
Beide sind nicht korrekt, aber welche der beiden Regressionsgeraden beschreibt die Stichprobe besser? Geben Sie hier auch die **Rechnungen** an, die Sie benötigen, um die beiden Regressionsgeraden zu vergleichen. **(9)**

3. Ein Professor stellt fest, dass die Noten in Statistik in den letzten 10 Semestern annähernd normalverteilt sind mit Erwartungswert 2,91 und Varianz 1,73.

Im Sommersemester 2007 wird von 55 Teilnehmern ein Notendurchschnitt von 3,19 erzielt. Die Studenten beschweren sich über eine zu schwere Klausur. Der Professor spricht von einer zufälligen Abweichung.

a) Kann die Hypothese des Professors, dass der Erwartungswert gleich geblieben ist (oder sich gar verbessert hat) aufrecht erhalten werden? Oder gilt die Alternative der Studenten, dass der Erwartungswert höher als 2,91 liegt? Verwenden Sie als Irrtumswahrscheinlickeit $\alpha = 0{,}05$. (Siehe auch Tabelle auf Seite 1.) **(12)**

b) Bei welcher Zahl von Teilnehmern (gerundet auf ganze Teilnehmer) würde gerade die Grenze zwischen Annahme und Ablehnung der Hypothese erreicht, wenn alle anderen Werte gleich bleiben? **(7)**

Lösungen

1. W_1 = Wert des Barpreises nach einem Jahr durch Verzinsung mit 5%:
$W_1 = 500 \cdot 1,05 = 525$

Erwartungswert für den Wert der Ratenzahlungen nach einem Jahr W_{R1}, wobei die erste Rate verzinst wird und die zweite nur mit 90% gerechnet wird, entsprechend der Wahrscheinlichkeit, dass sie gezahlt wird: $E(W_{R1}) = r \cdot 1,05 + r \cdot 0,9 = 1,95 \cdot r$

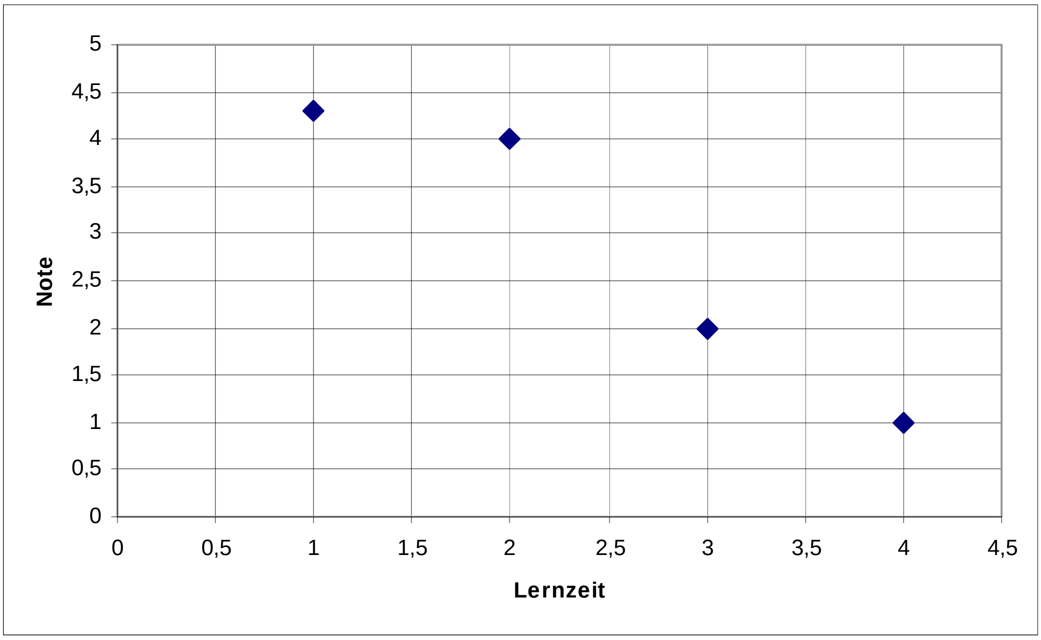
Beide Werte sollen gleich sein:

$$W_1 = E(W_{R1}) \Leftrightarrow 525 = 1,95 \cdot r \Leftrightarrow r = \frac{525}{1,95} = \underline{\underline{270,77}}$$

Die Rate r muss 270,77 € hoch sein, damit der Händler im Mittel bei Ratenzahlung nach einem Jahr die gleichen Einnahmen hat wie bei Barzahlung.

2. Stichprobe mit zwei Merkmalen

a) Streudiagramm



b) Wie groß ist der Korrelationskoeffizient? Sie brauchen ihn dazu nicht auszurechnen, sondern können ihn auch schätzen. Kreuzen Sie das richtige Kästchen an

☐ 0 ☐ 100 ☐ -1 ☐ +1 ☐ 1,5 ☐ -1,5 ☐ 95 ☐ -95

☐ 0,97 ☒ -0,97 ☐ 0,24 ☐ -0,24

c) Kreuzen Sie alle Begründungen an, die Ihre Wahl bei b) stützen. Setzen Sie maximal vier Kreuze. **(4)**

☐ Der Korrelationskoeffizient ist die Steigung der Regressionsgeraden

☐ Der Korrelationskoeffizient ist eine Zahl zwischen 0 und 1

☒ Der Korrelationskoeffizient ist eine Zahl zwischen -1 und +1

☐ Der Korrelationskoeffizient ist eine Zahl zwischen 0 und 100

☐ Der Korrelationskoeffizient ist eine Zahl zwischen -100 und +100

☒ Der Korrelationskoeffizient liegt nahe bei +1 oder -1, wenn die beiden Merkmale stark voneinander abhängen.

☒ Der Korrelationskoeffizient ist negativ, wenn die Steigung der Regressionsgeraden negativ ist.

☐ Der Korrelationskoeffizient kann nur dann 100 sein, wenn alle Punkte der Stichprobe exakt auf der Regressionsgeraden liegen.

☒ Der Korrelationskoeffizient kann nur dann 1 sein, wenn alle Punkte der Stichprobe exakt auf der Regressionsgeraden liegen.

☐ Der Korrelationskoeffizient ist 0, wenn alle Punkte der Stichprobe exakt auf der Regressionsgeraden liegen.

☐ Der Korrelationskoeffizient kann beliebig groß werden, wenn die Punkte weit verstreut liegen, da er die Summe der Abweichungs-Quadrate von der Regressionsgeraden ist.

d) Die Regressionsgerade beschreibt die Stichprobe dann am besten, wenn die Summe der Quadrate der Abweichung von berechneten Werten auf der Gerade zu den gemessenen Werten der Stichprobe minimal wird.

Schüler -Nr.	Lern zeit [h]	Note	berechnet Hans	Abweichungs Quadrat Hans	berech- net Thomas	Abweichungs Quadrat Thom
1	1	4,3	7-1,5=5,5	$1{,}2^2=1{,}44$	5-1=4,0	$0{,}3^2=0{,}09$
2	3	2,0	2,5	0,25	2,0	0
3	4	1,0	1,0	0	1,0	0
4	2	4,0	4,0	0	3,0	1
			••	••	•	•

Die Summe der Abweichungsquadrate bei Hans ist 1,69• und bei Thomas 1,09•. Die Gerade von Thomas ist daher besser. •

3. Gauß-Test für Mittelwert der Notenverteilung.

a) Umrechnung des beobachteten Mittelwerts auf Standardnormalverteilung

$$T(\overline{x}) = \frac{\sqrt{n} \bullet\bullet}{\sigma \bullet}(\overline{x} - \mu)\bullet = \frac{\sqrt{55}}{\sqrt{1,73}}(\overline{x} - 2,91) \bullet$$

$$T(3{,}19) = \frac{\sqrt{55}}{\sqrt{1,73}}(3{,}19 - 2{,}91) = 1,58 \bullet$$

Bestimmung der Grenze g: Es handelt sich um einen einseitigen Test, da nur dann die Hypothese in Frage gestellt wird, wenn der Notendurchschnitt schlechter wird. g ist damit eine obere Grenze.

$\Phi(g, \mu = 0, \sigma = 1)\bullet = 1 - \alpha = 0,95 \bullet$

$g = \Phi^{-1}(0,95; \mu = 0; \sigma = 1) = 1,64485 \bullet$ (aus Tabelle)

Vergleich von T(3,19) mit g: $T(3{,}19) = 1,58 < \bullet\bullet 1,64485 = g$

Der beobachtete Notendurchschnitt liegt daher im Annahmebereich des Tests, die Hypothese des Professors kann aufrecht erhalten werden. •

b) Wegen des zentralen Grenzwertsatzes geht die Wurzel aus der Teilnehmerzahl in den Test ein und zwar gilt für die Standardabweichung des Notendurchschnitts

$\sigma_{Notendurchschnitt} = \dfrac{\sigma_{Note}}{\sqrt{n}}$. Dies sieht man bei der Umrechnung auf T.

$$T(3{,}19) = \frac{\sqrt{n}}{\sigma}(\overline{x} - \mu) = \frac{\sqrt{n}}{\sqrt{1{,}73}}(3{,}19 - 2{,}91) = \frac{\sqrt{n} \cdot 0{,}28}{\sqrt{1{,}73}} = g$$

Diese Gleichung wird nach n aufgelöst:

$$\sqrt{n} = \frac{\sqrt{1{,}73} \cdot g}{0{,}28} \Rightarrow n = \frac{1{,}73 \cdot 1{,}64485^2}{0{,}28^2} = 59{,}70$$

Bei 59 Teilnehmern ist die Grenze g also gerade noch nicht überschritten, bei 60 Teilnehmern wäre man dagegen bereits im kritischen Bereich.