Klausur Statistik 2 RE
Statistik für Soziologen
Do, 24. 9. 2009

Name ……………………………….**Vorname**……………………….

Matrikelnummer ……………………………….

Einsichtnahme: Fr, 2. Oktober

BITTE <u>DEUTLICH UND LESERLICH</u> SCHREIBEN!

Es wird nur gewertet, was in diesem Exemplar steht. *Exemplar nicht zerlegen!*
Rechengänge sind nachvollziehbar anzugeben, Antworten zu begründen.

<u>Die gestellten FRAGEN sind zu BEANTWORTEN!</u>

1……….
2……….
3……….
4……….
5……….
6……….

Summe………….
Note……………..

1. In der folgenden Tabelle sind zu elf Objekten je drei Zahlen in Form von drei Datenspalten gegeben:

Objekt Nr	x	y	z
1	9	401	31
2	12	304	30
.	13	386	29
.	14	245	26
.	15	288	25
.	17	510	23
.	18	296	22
.	19	256	20
.	21	412	19
.	23	344	18
11	26	324	7

Zu diesen Daten liegen folgende sieben Zahlen vor:

-0,962 510 17 5649,9 -1 324 1,16

Welche dieser Zahlen ist der /die

Mittelwert der x	17
Pearson: korr(x, z)	-0.962
Rangkorrelation rs(x, z)	-1
Varianz(y)	5649.9
Median(y)	324

Zahl an der richtigen Stelle eintragen! **Zwei der Zahlen sind unbrauchbar**!

10

2. a. Ein ***fairer Würfel wird dreimal geworfen***. (Jede Augenzahl ist gleich wahrscheinlich)
Mit welcher Wahrscheinlichkeit erhält man bei dreimaligem Werfen dieses Würfels

- **höchstens zweimal die Sechs ?**
- **niemals eine Vier**
- **mindestens eine Drei**

b. Eine Zufallsgröße X sei ***normalverteilt N(160, 18)***

b1. Wie groß ist die Wahrscheinlichkeit **P(X = 169)**

b2. Bestimmen sie die **Wahrscheinlichkeit** $\mathbf{P(X \in [106, \ 178])}$

b3. Bestimmen sie **jene Zahl c** für die gilt: **P(X > c) = 71 Prozent**

5 + (1 + 2 + 2)

Lösung: a. ***215/216*** *= 0.995;* ***125/216 =*** *0.579* ***1 - P(keine Drei) = 91/216 =*** *0.421*

b1 ***0***
b2 ***0.8400***
b3 ***ca. 150.1***

3. Betrachtet werden drei Zufallsgrößen mit den angegebenen Verteilungen:

X : binomialverteilt mit den Parametern $n = 22$, $p = 0.1$
Y : stetig gleichverteilt auf $I = [1, 6]$
U: normalverteilt $N(3, 1.8)$

a. Bestimmen Sie jeweils die **Wahrscheinlichkeit** des **<u>offenen</u> Intervalls** $]\,0, 3[$
b. Wie groß ist bei jeder dieser Verteilungen der **Wert der Verteilungsfunktion** an der **Stelle 1.7** ?
c. Welcher Verteilung unterliegt ein **Mittelwert** von 100 genau wie U verteilten unabhängigen Zufallsgrößen und **mit welcher Wahrscheinlichkeit ist dieser Mittelwert kleiner als 3** ?

4 + 4+ 2

a. X: $P(\{1, 2\}) = 0.24072 + 0.28084 =$ **0.52156**

Y: $P(]\,0, 3[) = 2*0.2 =$ **0.4**

Z: $P(]\,0, 3[) = \Phi(0) - \Phi((0-3)/1.8) = 0.5 - \Phi(-1.67) = 0.5 – 0.0475 =$ **0.4525**

b. $FX(1.7) = P(\{0, 1\} = 0.09848 + 0.24072 =$ **0.3392**

$FY(1.7) = 0.7*0.2 =$ **0.14**

$FU(1,7) = \Phi((1.7-3)/1.8) = \Phi(-0.72) = 1- 0.7642 =$ **0.2358**

c. Verteilung von M_{100}: $N(3, 0.18)$ $P(M<3) =$ **0.5**

4. Eine Zufallsgröße unterliege einer $N(\mu, 7)$- Verteilung. Eine Stichprobe vom Umfang $n = 40$ ergab den Mittelwert 98.262.

a. **Lässt sich damit „bestätigen", dass der Erwartungswert unter 100 liegt?**
Testen Sie dazu die geeignete Gegenhypothese zum Niveau $\alpha = 0.1$ auf Signifikanz und interpretieren Sie das Ergebnis in Form einer **Antwort**!

b. Zu **welchem Signifikanzniveau** gehört der realisierte Mittelwert von 98.27 und was bedeutet das?

6 + 4

a.

$H_0 : \mu \geq 100 \qquad H_1 : \mu < 100$

Testgröße $t = \dfrac{98.262 - 100}{7} \cdot \sqrt{40} = -1.570$

Kritischer Bereich: $K =]-\infty, -\lambda_{0.1}] =]-\infty, -1.2816]$

t liegt in K also ist die Hypothese **„ μ < 100"** (Mit erlaubter maximaler Irrtumswahrscheinlichkeit 10 Prozent) bestätigt!

b.

Man bestimme α so, dass $t = -\lambda_\alpha \qquad \lambda_\alpha = 1.57$

$1.57 = \Phi^{-1}(1-\alpha) \qquad \Phi(1.57) = 1 - \alpha$

$1 - \alpha = 0.9418$ und daraus: der Signifikanzlevel $\alpha = 0.0582$ (= Irrtumswahrscheinlichkeit, wenn man aus den vorliegenden Daten schließt, der Erwartungswert liege unter 100)

5. In der Abluft zweier Entschwefelungsanlagen A und B wurde der verbliebene Schwefel in Mikrogramm je Kubikmeter an 40 Tagen (A) bzw. an 45 Tagen (B) gemessen und daraus wurden berechnet:

A: Mittelwert 51, und die Stichprobenstandardabweichung 12

B: Mittelwert 44, s = 16.

a. Kann aus den Daten mit einer Sicherheit von 95 % geschlossen werden, dass in B der gültige Grenzwert von 50 µg/m³ unterschritten wird?
Geht das auch zum Testniveau 0.001?

b. Lässt sich mit 95%iger Sicherheit schließen, dass B signifikant niedrigeren Schwefelgehalt aufweist als A?

Formulieren Sie immer genau die Null- und Alternativhypothese, geben Sie an, welcher Test durchzuführen ist, ob der Test einseitig oder zweiseitig ist und führen Sie ihn durch.
Setzen Sie dabei Normalverteilung und dort, wo es nötig ist, Varianzhomogenität voraus!

5 + 5

a. Einproben-t-Test

H_1: $\mu_B < 50$ $\quad t_0 = \frac{44-50}{16}\sqrt{45} = -2.516$

Testniveau 0.05: K =] $-\infty$, $-$ 1.68]

H1 signifikant; **Ja**, Grenzwert wird unterschritten

Testniveau 0.001: Tabellenwert größer als 2.678

H1 nicht hochsignifikant; **Nein**

b. Zweiproben-t Test H_1: Gehalt in B ist kleiner als in A

$H_1 : \mu_A > \mu_B$

Testgröße $t_0 = \frac{51-44}{s}\sqrt{4.602} = 2.2588$ s = 14.26

K = [1.664, ∞ [H_1 signifikant; **JA**

6. Es wird an mehreren zufällig ausgewählten Personen mit zwei unterschiedlichen Geräten jeweils der Blutdruck gemessen. Man erhielt:

Person	A	B	C	D	E	F	G	H .
Gerät XX	91	138	89	130	152	121	102	162
Gerät YY	93	144	90	129	150	128	104	163

a. Testen Sie zum Niveau $\alpha = 0.1$, ***ohne Annahme von Normalverteilungen***, die Nullhypothese „Die beiden Geräte liefern die selben Ergebnisse“

c. Testen Sie zum Niveau $\alpha = 0.1$, ***ohne Annahme von Normalverteilungen*** Sind die Messwerte von Gerät XX signifikant niedriger? (gleiches $\alpha = 0.1$)

(**Hinweis** für beide Aufgabenteile: trotz weniger Daten Normalverteilungstabellen anwenden!)

5 + 5

Abhängige Proben; nichtparametrisch, Lagevergleich:
Vorzeichenrangtest (Wilcoxon)

	A	B	C	D	E	F	G	H
x	91	138	89	130	152	121	102	162
y	93	144	90	129	150	128	104	163
Differenz: d = y – x	**2**	**6**	**1**	*-1*	*-2*	**7**	**2**	**1**
Rang von Betrag d:	**5**	**7**	**2**	*2*	*5*	**8**	**5**	**2**

Rangsumme r_+ **= 29**

E(T) = n*(n+1)/4 = 18 Var(T) = n*(n+1)*(2n+1)/24 = 51

$$t_0 = \frac{29-18}{\sqrt{51}} = 1.54$$

a. zweiseitig, $u_{0.95} = 1.6449$ t_0 nicht kritisch, Nullhypothese beibehalten, Gleichheit akzeptiert

b. einseitig, $u_{0.9} = 1.2816$
$K = [1.2816, \infty]$ t_0 kritisch, Geschwindigkeit am Vorderreifen ist signifikant höher