|  |  |
| --- | --- |
| HSB_Horizontal_RGB WS 2017/18 | **Prof. Dr. Peter Schmidt****SIB**www.schmidt-bremen.de |

**Statistik EFA**
Montag, 5.2.18

1. Im **Multiple-Choice-Teil** (Aufg. 1)ergibt jede richtige Lösung (zu einer Aussage) 1,5 Punkte.
Es können hier also insgesamt 21(+9) Punkte erzielt werden.
2. Die Lösungen von **Fill-In-Aufgaben** (Aufg. 2 bis 6) sind in die dafür vorgesehenen Felder (Kästchen) einzutragen. Dies gilt für Skizzen, Grafiken und Texte analog.
**Die Lösungen gelten nur dann, wenn der vollständige Lösungsweg erkennbar ist !**
3. Es ist NUR in der gehefteten Klausur zu arbeiten. Als Konzeptseiten können die Rückseiten der Blätter benutzt werden. **Die Blätter müssen geheftet bleiben**. Bitte benutzen Sie *keinen* Mantelbogen.
4. Es können 90 Punkte\* erreicht werden. Die erreichbare Punktzahl der einzelnen Aufgaben ist angegeben. Da auch die Bearbeitungszeit **90 Minuten** beträgt, sind die Punktzahlen ein Anhalt für die sinnvolle Bearbeitungszeit der Aufgaben. Tipp: Bearbeiten Sie möglichst viele Aufgaben.
*\* + 9 Zusatzpunkte aus der MC-Aufgaben, die Gesamtpunktzahl kann jedoch 90 nicht überschreiten*
5. Zulässige Hilfsmittel:
 - Formelsammlung (P. Schmidt) *ohne* eigene Erläuterungen
 - Taschenrechner ohne Textverarbeitungsfunktion
Bitte schreiben Sie nicht mit roten Stift.

***Viel Erfolg !***

1. **Modiltelefone (auch Smartwatches u.a.) sind verboten.** Schon Bereithalten **am Platz** gilt als Täuschungsversuch  in eine Tasche und diese entfernt vom Sitzplatz deponieren**!**
2. Überprüfen Sie zu Beginn die Klausur auf Vollständigkeit (6 Aufgaben auf 8 Seiten) und füllen
 bitte vorab die unten stehenden Kästchen aus *(„Versuch“ = ich schreibe die Klausur zum ... Mal)*. ⮷

**Lösungshinweise**

|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| **Name**: |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  | 1. Versuch:  |  |
| **Vorname**: |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  | 2. Versuch: |  |
| **MatrikelNr:** |  |  |  |  |  |  |  |  |  | Wenn zutreffend, **unbedingt** ankreuzen → **3.** Versuch: |  |

*Aufg. Punkte:*  ***Ab hier bitte nichts beschriften oder ankreuzen:***

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| 1.  | *(21+9)* |  |  |  |
| 2. | *(20)* |  | Datum:  |
| 3. | *(10)* |  |
| 4. | *(20)* |  | Prozent /  |  |
| 5. | *(10 +3)* |  | Note |  |
| 6. | *(9)* |  | Unterschrift |
| **** | ***(max. 90)*** |  |

**Achtung: Bitte keinen Leistungsnachweis-Bogen beilegen!**

Immer wieder 16.15 … ist die Mensa zu, genau dann wenn die kreative Phase vom kleinen Hunger unterbrochen wird – und auch von Durst. Auf dieser Basis entstand sie: „Die Wirtschaft“, in der Studierende der Werderstraße im 4. Stock des A-Gebäudes von 16:16-21:59 Uhr Heiß- und Kaltgetränke sowie kleine Snacks anbieten. *(Zu schön um wahr zu sein?)*

# **Multiple Choice** Bitte kreuzen Sie an - Erläuterungen sind nicht erforderlich. Hinweis: Eine Multiple-Choice-Aussage ist nur „richtig“, wenn die Aussage *immer* gilt. Gibt es ein einziges Gegenbeispiel, so ist sie „falsch“. *[je 1,5 Punkte ⇒ Gesamt 21+9 Punkte]*

|  |  |  |
| --- | --- | --- |
| richtig | falsch |  Welche Aussage ist richtig? |
|  | x | 1. Der Zentrale Grenzwertsatz beschreibt die Standardabweichungen der Mittelwerte
 |
| x |  | 1. Die induktive Statistik beschreibt das Verhältnis von Stichprobe und Population
 |
|  | x | 1. Mit einer Statistik des Brötchenkonsums kann man den Hunger der Studierenden beweisen.
 |
| x |  | 1. Ein Korrelationskoeffizient kann negative Werte annehmen.
 |
| x |  | 1. Ein Zweistichproben-Test untersucht z.B., ob zwei Kundengruppen (=Stichproben) aus der selben Grundgesamtheit stammen können.
 |
| x |  | 1. Nominal skalierte Daten sind in der Regel diskret.
 |
|  | x | 1. Das Bestimmtheitsmaß hat dasselbe Vorzeichen wie die Steigung b.
 |
| x |  | 1. Die Anzahl der möglichen Anordnungen (Reihenfolgen) von Studis an der Theke bezeichnet man als Permutationen.
 |
|  | x | 1. Lagemaße sind z.B. Mittelwerte, DAA, Quantile …
 |
| x |  | 1. n! = 1 \* 2 \* 3 \* ... \* n
 |
| x |  | 1. Ein Preisindex ist ein gewichteter Mittelwert aus Preissteigerungsziffern.
 |
|  | x | 1. Aus den Werten einer Ratingskala darf kein arithmetisches Mittel berechnet werden
 |
|  | x | 1. Wenn alle Residuen gleich 1 sind, wird das Bestimmtheitsmaß gleich 0.
 |
| x |  | 1. Durchschnittsnoten werden in Deutschland in der Regel als arithmetisches Mittel gebildet.
 |
|  | x | 1. Die Ermittlung von Durchschnittsnoten als arithmetisches Mittel ist statistisch korrekt.
 |
|  | x | 1. Ein Kaffee-Preisindex von 115 für 2018 zur Basis 2015 bedeutet, dass die Preise im Durchschnitt um 5 % pro Jahr gestiegen sind.
 |
|  | x | 1. Das LKM misst den Prozentsatz der maximalen Gleichheit.
 |
| x |  | 1. Ein Konfidenzintervall ist umso schmaler, je größer  ist.
 |  |
|  | x | 1. W(A ∪ B) = W(A) + W(B)
 |  |
| x |  | 1. Ein Konfidenzintervall ist umso schmaler, je größer n ist.
 | *(21+9)* |

# Sie analysieren Sie die Umsatzzahlen der StudierBar aus den letzten 3 Jahren. Es wurden jeweils Dritteljahre (Tertiale) aufgezeichnet. Sie sehen, dass die Entwicklung einer saisonalen Entwicklung folgt und möchten eine **saisonbereinigte Umsatzentwicklung** betrachten können. Für die Glättung verwenden Sie eine lineare Trendfunktion. Zusätzlich möchten Sie die zu erwartende Entwicklung für 2018 **prognostizieren.** *(Die Umsatzzahlen sind in TEuro angegeben, aber das brauchen Sie bei der Rechnung nicht zu berücksichtigen. Es ist ausreichend, wenn Sie mit einer Nachkommastelle rechnen)*  *[Gesamt: 20 Punkte]*



Hinweise: Bei Verwendung von einfachen t-Werten (t=1,2,3, ...) ergibt sich **a = 12,7**

## Beschreiben Sie die Entwicklung der Umsatzzahlen in den drei Jahren. Bewerten Sie diese Entwicklung. [5 Punkte]

s.o.: y^ = 20 + 1,47 \* ti\*  [bzw. y^ = 12,7 + 1,47 \* ti]

Positiver Trend, mehr Umsatz, gute Entwicklung

## Zeichnen Sie die Beobachtungswerte, die Gerade und die Werte aus Aufgabe 2.4 [4 Pkt]

 *(Zeichnung s.o.)*

## Ermitteln Sie auf dieser Basis die durchschnittlichen Saisonkomponenten für die Tertiale. Interpretieren Sie die „Saisonfigur“; in welchem Tertial wird am wenigsten verkauft? Haben Sie eine Idee, warum? [5 Punkte]

s.o. 2. Tertial – Sommerpause…, keine Studis da ☺

## Ermitteln Sie die saisonbereinigte Zeitreihe. Welchen Sinn hat eine solche Saisonbereinigung? [3 Punkte]

s.o. Wie der Name schon sagt, das Herausrechnen der immer wiederkehrenden saisonalen Einflüsse, damit die wirkliche (reale) Entwicklung des Marktes sichtbar wird.

## Ermitteln Sie die zu erwartenden Eintrittszahlen für die drei Tertiale 2016. Berücksichtigen Sie dabei die Saisoneinflüsse. Erläutern Sie Ihr Vorgehen in Stichworten. [3 Punkte]

s.o.

Es ist zu beachten, dass die oben ermittelten durchschnittlichen SK *addiert* werden, da Sie den reinen Trend überlagern. Nur so kann eine realistische Vorhersage der Entwicklung getroffen werden.

## Zeichnen Sie die Beobachtungswerte, die Gerade und die Werte aus 2.4 [4 Punkte]

s.o.

*(20)*

# Die Studentin Gamma arbeitet häufig in der StudierBar. Sie kennt inzwischen die Kunden recht gut. Grade kommen Peter S., kurz danach Klaus F. Sie überlegen immer lange, ob sie ein Bier trinken wollen. Letztlich tun sie das beide jeweils in 9 von 10 Fällen. [10 Punkte]

## Wie hoch ist die Wahrscheinlichkeit, dass ... [4 Punkte]

|  |  |
| --- | --- |
| beide ein Bier bestellen? | W(A ∩ B) = 0,9 \* 0,9 = 0,81 oder 81 Prozent |
| genau EineR von beiden ein Bier nimmt? | W((A ∩ B-)∪ (A -∩ B))= 0,9 \* 0,1 + 0,1 \* 0,9 =0,18 od. 18% |
| Peter eins trinkt, Dorothea aber nicht | W(A ∩ B-) = 0,9 \* 0,1 = 0,09 oder 9 % |
| keiner ein Bier bestellt? | W(A- ∩ B-) = 0,1 \* 0,1 = 0,01 oder 1 % |

## Zum Getränk gibt es verschiedene Sorten von Chips. 3 Marken sind mild und 2 sind von kräftigem Geschmack. Wie groß ist die Wahrscheinlichkeit, dass Peter, der sich ohne hinzusehen 2 greift (ohne Zurücklegen natürlich), zwei kräftige erwischt? [2 Punkte]

 2 / 5 \* 1 / 4 = 2 / 20 = 10 Prozent

## Bei einer Werbeaktion dürfen die Kundinnen und Kunden würfeln. Für jede 6 gibt es einen leckeren Käsehappen. Dorothea würfelt dreimal. Wie groß ist die Wahrscheinlichkeit, dass sie) genau einmal b) genau zweimal c) genau dreimal d) mindestens zweimal zubeißen darf. (d.h. jeweils dies gefragte Anzahl 6en würfelt)? [4 Punkte]

 *a) 1/6 \* 5/6 \* 5/6 \*3 = 25/72 oder 34,72 Prozent*

*b) 1/6 \* 1/6 \* 5/6 \*3 = 5/72 oder 6,94 Prozent*

*(10)*

*c) 1/6 \* 1/6 \*1/6 = 1/216 oder 0,46 Prozent*

*d) “b) + c)” = 15/216 + 1/216 = 2/27 oder 7,4 Prozent*

# Fett ... oder nicht fett ? Ein wesentliches Qualitätskriterium für den angebotenen Tortenbrie der Wahrheit ist der Fettgehalt. Um eine hochwertige Produktqualität sicherzustellen, achtet die StudierBar verstärkt auf dieses Kriterium.  *[Ges: 20 Pkt]*

## Für einen guten Tortenbrie ist es wichtig, dass der Fettgehalt von 20 % weder überschritten (aus Geschmacksgründen) noch unterschritten wird (zu fett). Die Lieferfirma NordKäs AG hat dies zugesichert. Berndt untersucht 250 Proben und ermittelt einen Fettgehalt von 23 %. Kann die StudierBar das Produkt gebrauchen, wenn eine Irrtumswahrscheinlichkeit von 5 % gilt? [7 Punkte]

|  |  |  |  |
| --- | --- | --- | --- |
| Informationen: | Beobachtungen | Mittelwert | Verteilung |
| Grundgesamtheit | k.A. | **po = 0,20** (Aufgabe) | Normalverteilung wird angenommen |
| Stichprobe | 250 | **= 0,23**  |

1) Hypothesen: Ho: p = p0 = 0,1 ; H1: p ≠ p0 → zweiseitiger Test

2)  = 0,05 lt. Aufgabenstellung

3) Keine Endlichkeitskorrektur (da kein N angegeben, als „sehr groß“ angenommen) →
 $σ\_{p\^}\sqrt{\frac{0,23\*0,77}{250}}$ = **0,0266**

4) zc = 1,96

~~5) z~~~~p~~ ~~= 1,41 [Alternative B]~~ **pcu = 0,148 pco = 0,252**

6) |zp| > |zc| ? Nein! → H0 NICHT verwerfen ** > pcu und < pco → H0 nicht verwerfen**

7) Der Käse kann tatsächlich aus einer Grundgesamtheit mit 20% Fettgehalt stammen
 und daher kann er verwendet werden.

## Schätzen Sie auf Basis o.a. Stichprobe, wie groß der Fettgehalt des Tortenbrie der Wahrheit in der Grundgesamtheit ist.  Was heißen diese Werte inhaltlich? (=auf Deutsch in Ihren Worten) Finden Sie die Schätzung plausibel? Wann wäre der Bereich kleiner? [3 Punkte]

 *Konfidenzintervall*  *W (0,23 –* 1,96*⋅* 0, 0266*≤ p ≤ 0,23 +* 1,96*⋅* 0, 0266*) = 0,95
 W (****0,178*** *≤  ≤* ***0,282****) = 0,95*

*Mit 95prozentiger Wahrscheinlichkeit liegt der Fettgehalt zwischen 17,8 und 28,2 Prozent. Dies ist eine Schätzung des unbekannten Anteilswertes der Grundgesamtheit auf Basis eines (zufälligen) Stichprobenanteils.*

*Das ist ein ziemlich breites Intervall und für die Entscheidung nicht besonders hilfreich.
Bei höherem  würde das Intervall kleiner, ebenso bei einer größeren Stichprobe.*

## Skizzieren Sie dieses Intervall zeichnerisch [3 Punkte]

 *f(**)*

 *gu  go*

 *0,178*  *0,282* 

## Bei den Erdnuss-Snacks ist es wichtig, dass mindestens 50 gr. in der Schale sind. Die StudierBar testet 144 Proben und kommt auf einen Durchschnittswert von 49,2 gr. bei einer Standardabweichung von 4 gr. Testen Sie, ob damit bei einem Sicherheitsgrad von 95 Prozent die Vorgabe eingehalten wird. [7 Punkte]

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Informationen | Beobachtungen | Mittelwert | Standardabweichung | Verteilung |
| Grundgesamtheit | k.A.  | o = 50 | k.A. | k.A. |
| Stichprobe | 144 | = 49,2 | 4 |  |

1) Hypothesen: Ho: μ > μo = 50 ⇒ H1: μ < μo → einseitig (links) kritisch

2)  = 0,05 lt. Aufgabenstellung

3)  unbekannt, n > 50 → 2. Fall
keine Endlichkeitskorrektur (da N als „groß“ angenommen wird) → **= 0,333**

4) zc = 1,64

5) zx = –2,4 [Alternative B] **cu = 49,45 co = 50,55**

6) |zx| > |zc| → Ho verwerfen ** < cu → Ho verwerfen**

7) In diesem Test wird die Aussage, dass die Anzahl groß genug ist, abgelehnt, d.h. also es sind zu wenig Erdnüsse in der Schale.

*(20)*

# Die StudierBar befragt zwei Kunden (Peter und Paul), welches Gebäck sie am liebsten mögen. *[Gesamt: 10 Punkte]*

|  |  |  |
| --- | --- | --- |
|  | Rangfolge  |  |
|  | Peter  | Paul |  |
| Brownie | 1 | 4 |  |
| Bluei | 2 | 5 |  |
| Yellowi | 3 | 2 |  |
| Möhrchen | 4 | 1 |  |
| Cookie  | 5 | 3 |  |
|  |  |  |  |

## Gibt es einen Zusammenhang zwischen den Bewertungen der Ernährungs-Expertinnen? Welches ist das geeignete Maß, die Stärke des Zusammenhanges zu ermitteln? [3 Punkte]

|  |
| --- |
| Geeignet ist der Rangkorrelationskoeffizient nach SpearmanGrund: ordinal skaliert |

## Errechnen Sie dieses Maß und interpretieren Sie das Ergebnis. [4 Punkte]

|  |
| --- |
| Mittlerer negativer Zusammenhang = die Geschmäcker der beiden sind unterschiedlich |

## Welche anderen Zusammenhangmaße gibt es? Beschreiben Sie diese [3 Punkte]

|  |
| --- |
| Hier sollen die verschiedenen Maße in Abhängigkeit von den jeweiligen Skalen beschrieben werden |

BONUS:

## Sie schreiben eine Hausarbeit über die StudierBar und sollen den Absatz der **Torten der Wahrheit** statistisch beschreiben. Welche statistischen Methoden könnten Sie verwenden. [3 Punkte]

|  |
| --- |
| … Bonus …  |

# Gleitende Durchschnitte [Gesamt: 9 Punkte]

## Für die (Zeitreihen-) Analyse der Anzahl der KundInnen um 21.00 Uhr an sechs aufeinander folgenden Tagen möchten Sie nun die folgenden Werte glätten. Verwenden Sie dazu Gleitende Durchschnitte dritter und vierter Ordnung[[1]](#footnote-1). [~~5~~ P]

 *(Zeichnung s.u.)*

## Zeichnen Sie die beiden gleitenden Durchschnitte in die u.a. Grafik ein und erläutern Sie kurz den „Informationsverlust am Rand“. [4 P]



*(9)*

BONUS:

## Kommentieren Sie die u.a. Überschrift aus dem Handelsblatt vom 31.1.18 aus statistischer Sicht. [3 Punkte]



|  |
| --- |
| … Bonus …  |

1. Es ist freigestellt, ob Sie zentrierte oder endwertige GD verwenden. [↑](#footnote-ref-1)