Fernstudium4You.de

Klausurlösung
Grundlagen der Statistik

WS2018

Vorwort

Dieses Dokument beinhaltet die Lösung der Klausuraufgaben vom Wintersemester 2018 für den Teil der **Grundlagen der Statistik** aus dem Modul der **Grundlagen der Wirtschaftsmathematik und Statistik** vom Lehrstuhl Prof. Dr. Singer von der Fernuni Hagen.

Aus urheberrechtlichen Gründen beinhaltet dieses Dokument nicht die Aufgabenstellungen. Die Originalklausuraufgaben können auf der Seite vom Lehrstuhl oder der Seite der Übungsklausuren heruntergeladen werden:

http://www.fernuni-hagen.de/ls_statistik/klausuren

https://www.fernuni-hagen.de/wirtschaftswissenschaft/studium/uebungsklausuren.shtml

Es sei darauf hingewiesen, dass es sich bei den Lösungen **nicht** um Musterlösungen der Fernuni Hagen handelt. Bei den Lösungswegen handelt es sich um Empfehlungen. Auch andere Vorgehensweise führen zur maximalen Punktzahl.

Klausurlösung Grundlagen der Statistik WS2018

(http://www.fernuni-hagen.de/wirtschaftswissenschaft/studium/download/pruefungen/31101_ws1718.pdf)

Lösung Aufgabe 7

Die Aussagen A)-E) behandeln die Grundlagen der Merkmale

A)

Richtig. Metrische Merkmale werden nochmal in die drei verschiedenen Skalen: Intervallskala, Verhältnisskala und Absolutskala unterschieden. Die Zuordnung der metrischen Merkmale in diese drei verschiedenen Skalen erfolgt aufgrund unterschiedlicher Größen, wobei mit Größen Maßeinheiten verschiedener Arten (Grad Celsius, Meter, etc.) gemeint ist. Da diese Größen alle in Form von Zahlenwerten erhoben werden können, sind diese quantitativ.

Die Aussage ist richtig.

B)

Falsch. Nominalskalierte Merkmale können nicht als Zahlenwerte erhoben werden, entsprechend liegen diese Merkmale qualitativ vor.

Die Aussage ist falsch.

C)

Richtig. Die dichotome Grundgesamtheit ist sozusagen die Binomialverteilung. Wie wir wissen, gibt es in der Binomialverteilung nur zwei mögliche Merkmalsausprägungen bzw. Möglichkeiten, die zur Auswahl stehen. In diesem Fall ist es „Eigenschaft vorhanden“ oder „Eigenschaft nicht vorhanden“.

Die Aussage ist richtig.

D)

Falsch. Misst man die Lebensdauer, Größe oder das Gewicht genügend genau, erheben sich für die erhobenen Zahlenwerte Stellen hinter dem Komma (4,5 Jahre, 1,86m, 70,5kg). Weisen erhobene Merkmalswerte Stellen hinter dem Komma auf, bzw. sind nicht ganzzahlig, werden diese als stetige Merkmale bezeichnet.

Die Aussage ist falsch.

E)

Richtig. Nominalskalierte Merkmale unterscheiden sich ausschließlich nur nach ihrer Gleichheit (Häufigkeiten bei Beobachtungen) oder Verschiedenheit (mehrere Ausprägungen), können jedoch nicht in eine Rangordnung gebracht werden.

Die Aussage ist richtig.

A, **C** und **E** sind richtig.

B und D sind falsch.

Lösung Aufgabe 8

Die Aussagen A)-E) behandeln die Thematik der Lage- und Streuungsmaße.

Vor dem Lösen dieser Aufgabe sind die angegebenen Merkmalswerte nach ihrer Reihenfolge zu sortieren:

1 1 2 3 3 7 7 7 8 10

A)

Der Modalwert entspricht der Merkmalsausprägung, die am häufigsten vorkommt:

$x_{mod} = \max h(x_j) = 7$

Der Merkmalswert 7 kommt am häufigsten vor.

Die Aussage ist richtig.

B)

Es existiert keine Merkmalsausprägung die ebenfalls mit der gleichen Häufigkeit wie die Merkmalsausprägung 7 vorkommt. Entsprechend gibt es nur ein Modalwert.

Die Aussage ist falsch.

C)

Die Anzahl der Merkmalswerte $n = 10$ ist mit 0,5 zu multiplizieren:

$x_{0,5} = n \cdot 0{,}5 = 10 \cdot 0{,}5 = 5$

Die Zahl ist gerade, also müssen wir den Durchschnitt aus der Merkmalsausprägung an 5ter und 6ter Stelle berechnen:

$$x_{med} = \frac{3+7}{2} = 5$$

Der Median beträgt 5.

Die Aussage ist falsch.

D)

Es ist das arithmetische Mittel zu berechnen:

$$\bar{x} = \frac{1+1+2+3+3+7+7+7+8+10}{10} = \frac{49}{10} = 4{,}9$$

Die Aussage ist richtig.

E)

Es ist zunächst die Varianz der Merkmalswerte zu berechnen:

$$\tilde{s}^2 = \frac{2\cdot(1-4{,}9)^2 + (2-4{,}9)^2 + 2\cdot(3-4{,}9)^2 + 3\cdot(7-4{,}9)^2 + (8-4{,}9)^2 + (10-4{,}9)^2}{10}$$

$$\tilde{s}^2 = \frac{94{,}9}{10} = 9{,}49$$

Nun gilt es die Beobachtungswerte mit 2 zu multiplizieren und das arithmetische Mittel erneut auszurechnen, wobei sich hier auch schnell erkennen lässt, dass sich das arithmetische Mittel ebenfalls verdoppelt. Die Reihe von Merkmalswerten verdoppelt lautet:

2 2 4 6 6 14 14 14 16 20

$$\bar{x} = \frac{2+2+4+6+6+14+14+14+16+20}{10} = \frac{98}{10} = 9{,}8$$

Es gilt erneut die Varianz zu berechnen:

$$\tilde{s}^2 = \frac{2\cdot(2-9{,}8)^2 + (4-9{,}8)^2 + 2\cdot(6-9{,}8)^2 + 3\cdot(14-9{,}8)^2 + (16-9{,}8)^2 + (20-9{,}8)^2}{10}$$

$$\tilde{s}^2 = \frac{379{,}6}{10} = 37{,}96$$

Die Varianz $\tilde{s}^2$ verändert sich um den Faktor $\frac{37{,}96}{9{,}49} = 4$.

Merke: Diese Aufgabe hätte man leichter mit den Rechenregeln für Varianzen lösen können, da wir wissen, dass ein Multiplikator/Faktor immer zu quadrieren ist, um zur Varianz zu gelangen:

$$\tilde{s}^2 = \mathbf{2^2} \cdot 9{,}49 = \mathbf{4} \cdot 9{,}49 = 37{,}96$$

Die Aussage ist falsch.

A und **D** sind richtig.

B, C und E sind falsch.

Lösung Aufgabe 9

Die Aussagen A) - E) behandeln die Thematik der Korrelationsrechnung.

A)

Falsch. Der Korrelationskoeffizient nach Bravais-Pearson kann lediglich für metrisch skalierte Merkmale berechnet werden. Ist ein Merkmal ordinalskaliert kann lediglich der Korrelationskoeffizient nach Spearman berechnet werden.

Die Aussage ist falsch.

B)

Richtig. Der Korrelationskoeffizient nach Spearman trifft eine Aussage über die Stärke des monotonen Zusammenhangs zwischen zwei Merkmalen.

Die Aussage ist richtig.

C)

Richtig. Da nach dem „betragsmäßigen" Zusammenhang gefragt wird, ist es unerheblich, ob der Korrelationskoeffizient den Wert 1 oder -1 annimmt, die Beobachtungswerte liegen in jedem Fall entweder auf einer steigenden oder fallenden Geraden.

Die Aussage ist richtig.

D)

Falsch. Korrelationskoeffizienten können lediglich eine Aussage über die Stärke des Zusammenhangs treffen, nicht jedoch über die Kausalität des Zusammenhangs.

Die Aussage ist falsch.

E)

Falsch. Nur die Unabhängigkeit von Merkmalen zeigt auf, dass kein Zusammenhang zwischen den Merkmalen existiert. Der Korrelationskoeffizient nach Bravais-Pearson trifft lediglich eine Aussage über den linearen Zusammenhang zwischen den Merkmalen.

Die Aussage ist falsch.

B und **C** sind richtig.

A, D und E sind falsch.

Lösung Aufgabe 10

Die Aussagen A) – E) behandeln die Thematik der Grundlagen der Wahrscheinlichkeitsrechnung.

A)

Falsch. Hier wurde der Additionssatz umgestellt:

$$P(X \cup Y) = P(X) + P(Y) - P(X \cap Y) = P(X) + P(Y) - P(Z)$$

Umformen:

$$P(X \cup Y) = P(X) + P(Y) - P(Z) \quad | + P(Z)$$

$$P(X \cup Y) + P(Z) = P(X) + P(Y)$$

Die Schnittmenge muss zur Vereinigungsmenge addiert anstatt subtrahiert werden.

Die Aussage ist falsch.

B)

Richtig. Wenn Z die Schnittmenge zwischen X und Y darstellt, kann die Schnittmenge Z kleiner oder gleich der Menge von Y sein. Die Gleichheit ist dann der Fall, wenn es sich bei der Menge von Y um eine Teilmenge von X handelt.

Die Aussage ist richtig.

C)

Richtig. Für $P(X|Y)$ gilt:

$$P(X|Y) = \frac{P(X \cap Y)}{P(Y)}$$

Da $P(Z)$ mit der Schnittmenge von X und Y übereinstimmt, kann man auch schreiben:

$$P(X|Y) = \frac{P(X \cap Y)}{P(Y)} = \frac{P(Z)}{P(Y)}$$

Multiplizieren wir diesen Term mit $P(Y)$ gelangen wir zu:

$$\frac{P(Z)}{P(Y)} \cdot P(Y) = P(Z)$$

Die Aussage ist richtig.

D)

Richtig. Hier wurde der Additionssatz umgestellt:

$$P(X \cup Y) = P(X) + P(Y) - P(X \cap Y) = P(X) + P(Y) - P(Z)$$

Umformen:

$$P(X \cup Y) = P(X) + P(Y) - P(Z) \quad | + P(Z) \quad | - P(X \cup Y)$$

$$P(Z) = P(X) + P(Y) - P(X \cup Y)$$

Die Aussage ist richtig.

E)

Entfällt.

Die Aussage ist falsch.

B, C und **D** sind richtig.

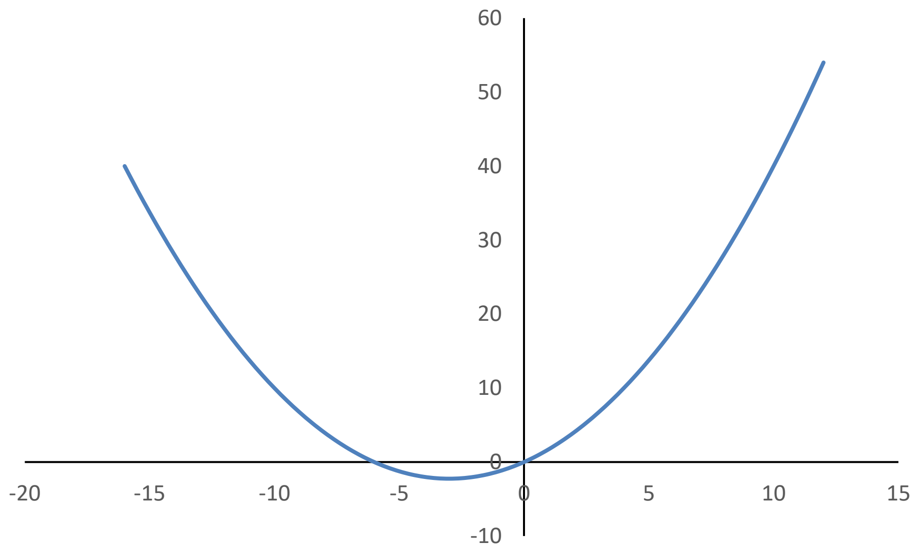
A und E sind falsch.

Lösung Aufgabe 11

Die Aussagen A) – E) behandeln die Thematik der Regressionsanalyse.

A)

Falsch. Bei der angegebenen Funktion handelt es sich um eine Parabel. Der Regressionskoeffizient a beschreibt bei der Parabel lediglich den Wert, wo die Parabel die y-Achse schneidet, nämlich bei 0:



Die Aussage ist falsch.

B)

Falsch. Bei dem angegebenen Funktionstyp handelt es sich um eine Parabel.

Die Aussage ist falsch.

C)

Richtig. Zur Berechnung des Prognosewertes ist $x = 2$ in die angegebene Funktion einzusetzen, wobei die Regressionskoeffizienten a, b und c mit ihren angegebenen Werten in der Aufgabenstellung zu ersetzen sind:

$\hat{y} = a + bx + cx^2$

$\hat{y} = 0 + 1{,}5x + 0{,}25x^2$

Einsetzen von $x = 2$:

$\hat{y} = 0 + 1{,}5 \cdot 2 + 0{,}25 \cdot 2^2 = 4$

Die Aussage ist richtig.

D)

Falsch. Regressionsfunktionen können beliebige Funktionen bzw. Formen annehmen, je nachdem, welcher Zusammenhang zwischen den Merkmalen vermutet wird. So kommt als Regressionsfunktion beispielsweise auch die Linearfunktion, die Potenzfunktion oder auch die Exponentialfunktion in Frage.

Die Aussage ist falsch.

E)

Entfällt.

Die Aussage ist falsch.

C ist richtig. A, B, D und E sind falsch.

Lösung Aufgabe 12

Die Aussagen A) - C) behandeln die Thematik der Parameter von Wahrscheinlichkeitsverteilungen. Aussagen D) und E) hingegen die speziellen Wahrscheinlichkeitsverteilungen.

A)

Zur Berechnung des Erwartungswertes sind die angegebenen Wahrscheinlichkeiten mit den möglichen Realisationen zu multiplizieren.

$$E(X) = \frac{1}{4} \cdot (-2) + \frac{1}{4} \cdot 0 + \frac{1}{4} \cdot 1 + \frac{1}{4} \cdot 3 = \frac{1}{2} = 0{,}5$$

Die Aussage ist falsch.

B)

Richtig. Siehe die Berechnung in A).

Die Aussage ist richtig.

C)

Falsch. Siehe die Berechnung in C)

Die Aussage ist falsch.

D)

Richtig. Da jede Realisation mit der gleichen Wahrscheinlichkeit eintritt, handelt es sich um eine Gleichverteilung.

Die Aussage ist richtig.

E)

Falsch. Eine Binomialverteilung ist auf zwei mögliche Realisationen begrenzt (Ereignis tritt ein oder nicht). Bei der Binomialverteilung beschreibt n auch nicht die Anzahl der Möglichen Realisationen, sondern die Anzahl der Wiederholungen des Zufallsexperiments.

Die Aussage ist falsch.

B und **D** sind richtig. B, C und E sind falsch.

Lösung Aufgabe 13

Die Aussagen A)-E) behandeln die Grundlagen der statistischen Testverfahren.

A)

Richtig. Da nur die **Zunahme** der Luftverschmutzung im Mittelpunkt steht und eine mögliche Abnahme nicht betrachtet werden soll, sollte ein einseitiges Testverfahren durchgeführt werden.

Die Aussage ist richtig.

B)

Falsch. Da der Kolbendurchmesser von der Norm sowohl nach oben als auch nach unten abweichen kann, ist hier ein zweiseitiges Testverfahren zu verwenden.

Die Aussage ist falsch.

C)

Richtig. Da nur das **Überwinden** der 5% Hürde im Mittelpunkt steht, sollte ein einseitiges Testverfahren durchgeführt werden.

Die Aussage ist richtig.

D)

Falsch. Eine Veränderung des Energieverbrauchs der Bevölkerung pro Kopf und Jahr ergibt sich, wenn entweder mehr oder weniger Energie verbraucht wird. Entsprechend sollte ein zweiseitiges Testverfahren verwendet werden.

Die Aussage ist falsch.

E)

Richtig. Da nur ein **Abfall** der Nachfrage eines Gutes im Mittelpunkt steht und eine Zunahme nicht in Betracht gezogen werden soll, sollte ein einseitiges Testverfahren durchgeführt werden.

Die Aussage ist richtig.

A, **C** und **E** sind richtig.

B und D sind falsch.

Lösung Aufgabe 14

Die Aussagen A) – E) behandeln die Thematik nicht-parametrischen statistischen Testverfahren (Teil 2).

A)

Richtig. Es handelt sich hierbei um einen Vorzeichentest. Bei diesem ist der Anteilswert π immer mit 0,5 zu wählen und n in Höhe des Stichprobenumfangs abzüglich der Stichproben, die als Differenz gleich sind, was hier allerdings nicht der Fall ist. So ergibt sich für die Testgröße die Binomialverteilung

$B(n, \pi) = (20\,; 0{,}5)$

Die Aussage ist richtig.

B)

Da zwei Grenzen zu ermitteln sind, sind in der Binomialverteilung die x-Werte zu suchen, bei dem die Wahrscheinlichkeit der Verteilungsfunktion $F_x(x)$ zum ersten Mal $\frac{\alpha}{2}$ überschreitet, für die untere Grenze, und zum ersten Mal $1 - \frac{\alpha}{2}$ überschreitet oder gleich ist, für die obere Grenze.

$$F_X(c_u) > \frac{0{,}1}{2} = F_X(c_u) > 0{,}05 = 6$$

$$F_X(c_o) \geq 1 - \frac{0{,}1}{2} = F_X(c_o) \geq 0{,}95 = 14$$

Die Grenzen lauten $c_u = 6$ und $c_o = 14$

Die Aussage ist richtig.

C)

Falsch. Siehe B).

Die Aussage ist falsch.

D)

Zur Ermittlung von Z_n ist die Anzahl der Plus-Zeichen (+) zu zählen:

$Z_n = 5$

Die Aussage ist richtig.

E)

Da gilt, dass $Z_n = 5 < c_u = 6 < c_o = 14$, kann die Nullhypothese abgelehnt, bzw. verworfen werden.

Die Aussage ist falsch.

A, B und **D** sind richtig. C und E sind falsch.

Lösung Aufgabe 45 (numerisch)

Diese Aufgabe behandelt die Thematik der Normalverteilung.

Zur Bestimmung der Wahrscheinlichkeit P ist die gemeinsame Verteilung der normalverteilten Zufallsvariablen zu bestimmen. Dazu sind die Zufallsvariablen B_1 , B_2 und B_3 zu addieren:

$$B_1 \sim N(1200\,;10092) + B_2 \sim N(1200\,;10092) + B_3 \sim N(1200\,;10092) = B \sim N(3600\,;30276)$$

Es ist die Wahrscheinlichkeit dafür herauszufinden, dass die Brenndauer zwischen 3200 und 3800 Stunden liegt. Folglich gilt für die normalverteilte Zufallsvariable im angegebenen Intervall:

$$P(3200 \leq X \leq 3800)$$

Zur Bestimmung der Wahrscheinlichkeit sind die angegebenen Grenzen durch eine lineare Transformation in die Standardnormalverteilung zu überführen, damit sich die Wahrscheinlichkeit für das Intervall über die Standardnormalverteilung ermitteln lässt:

$$P\left(\frac{3200 - 3600}{\sqrt{30276}} \leq Z \leq \frac{3800 - 3600}{\sqrt{30276}}\right) = P(-2{,}30 \leq Z \leq 1{,}15) = 0{,}8749 - (1 - 0{,}9893)$$

$$P(-2{,}30 \leq Z \leq 1{,}15) = 0{,}8642$$

Die Wahrscheinlichkeit, dass die Brenndauer zwischen 3200 und 3800 Stunden liegt beträgt 0,8642.

Die Lösung lautet 0,8642.

Lösung Aufgabe 46 (numerisch)

Diese Aufgabe behandelt die Thematik der Parameter von Wahrscheinlichkeitsverteilungen.

Zur Ermittlung des Quantil-Wertes einer stetigen Verteilungsfunktion, ist die Verteilungsfunktion gleich dem Quantil-Wert (0,5) zu setzten.

ACHTUNG: Bevor man den Wert gleich der Verteilungsfunktion setzt, sollte die Verteilungsfunktion ausmultipliziert werden:

$$F_X(x) = \frac{1}{100}(x-2)^2 = \frac{1}{100} \cdot (x^2 - 4x + 4) = \frac{1}{100}x^2 - 0{,}04x + 0{,}04$$

Die Verteilungsfunktion ist nun gleichzusetzen und nach Null umzustellen, sodass die PQ-Formel Anwendung finden kann:

$$\frac{1}{100}x^2 - 0{,}04x + 0{,}04 = 0{,}5 \quad | - 0{,}5$$

$$\frac{1}{100}x^2 - 0{,}04x - 0{,}46 = 0 \qquad |: \frac{1}{100}$$

$$x^2 - 4x - 46 = 0$$

PQ-Formel:

$$x_{1,2} = -\left(-\frac{4}{2}\right) \pm \sqrt{\left(\frac{4}{2}\right)^2 - (-46)} = 2 \pm \sqrt{2^2 + 46} = 2 \pm 7{,}07$$

$$x_{med} = 2 + 7{,}07 = 9{,}07 \approx 9{,}1$$

Da nur positive Werte als Quantil-Werte in Frage kommen, da die Verteilungsfunktion einen Gültigkeitsbereich von $2 \leq x < 12$ hat, ermittelt sich der Median eindeutig zu 9,1.

Die Lösung lautet 9,1.

Lust auf weitere Klausurlösungen?

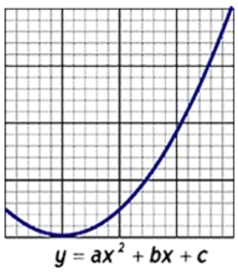
Das Mathe-Paket II enthält die **Klausurlösungen** des Statistik-Teils vom Wintersemester 2013 bis zum Wintersemester 2018 (**WS2013-WS2018**) und vieles mehr:

Videovortrag

24 Stunden qualitativ hochwertige, strukturierte und motivierende Videovorträge zum Thema Statistik. Der Videovortrag hilft dir viele Thematiken zu verstehen, ohne das Skript zuvor gelesen zu haben. Der Vortrag ist auf allen gängigen Geräten verfügbar, sowohl auf dem PC, Tablet als auch Smartphone.

Folienskript

Auf 240 Seiten erhältst du alle Folien der Präsentation zur visuellen Unterstützung und zum besseren Verständnis. Nach dem Ausdrucken kannst du hier Notizen und für dich sinnvolle Ergänzungen während des Videovortrags machen, als auch gänzlich ohne Internet mit dem Mathe-Paket II arbeiten und lernen



Formelsammlung

Die Formelsammlung enthält alle relevanten Formeln. Auf einem Blick hast du eine sofortige Übersicht der Formeln, als auch eine Beschreibung der einzelnen Symbole an der Seite. So kannst du während des Lernens schnell nachschlagen und musst nicht seitenweise durch das Skript blättern. Zusätzlich kannst du damit schnell und einfach die wichtigsten Formeln im Glossar markieren.

Übungsaufgaben

Zu jeder Thematik des Videovortrages gibt es passende Übungsaufgaben im Multiple-Choice Stil, angelehnt an die Klausuraufgaben. Insgesamt stehen dir über 160 Übungsaufgaben zur Verfügung, sodass du das Gelernte in die Praxis umsetzten kannst und damit dein Wissen weiter vertiefst.

Übungsklausuren

Dir steht eine unbegrenzte Anzahl an Übungsklausuren zur Verfügung. Aus allen über 160 Übungsaufgaben werden wahllos genau 10 Aufgaben ausgewählt, die eine Klausur darstellen. Nach Bearbeitung der Aufgaben erhältst du eine persönliche Klausurnote und detailliertes Feedback, was richtig und was falsch war. Durch das Simulieren einer Klausur kannst du dich perfekt auf den Klausurablauf vorbereiten.