

Statistik Vorlesung

05. Juli 2024

Dauer der Prüfung: 100 Minuten

ZUNAME:			
VORNAME:		MATR.NR.:	

ERLAUBT: **Skriptum des Instituts, nicht-graphikfähiger Taschenrechner**

VERBOTEN: **alle sonstigen Unterlagen, graphikfähiger Taschenrechner, Handys**

Bei den Single-Choice-Fragen bringt eine richtige Antwort 2 Punkte und eine falsche 1 Punkt Abzug. Es gibt keine negative Punktemitnahme in ein anderes Beispiel.

Berechnungen müssen nachvollziehbar aufgeschrieben sein.

Aufgabe	max. Punkte	erreichte Punkte
1	12	
2	10	
3	20	
4	20	
5	30	
6	8	
Summe	100	
Note		

1. (12 Punkte)

a) Der Umsatzindex kann auch negativ werden.

Richtig Falsch

b) Wenn der Pearson'sche Korrelationskoeffizient negativ ist, liegt die Steigung der Regressionsgeraden im Intervall $[-1; 0]$.

Richtig Falsch

c) Für eine Zufallsvariable X , die stetig gleichverteilt auf $[1,6]$ ist, gilt:

$$P(X = 1) = P(X = 2) = \dots = P(X = 6) = 1/6.$$

Richtig Falsch

d) Die Dichtefunktion jeder Normalverteilung ist symmetrisch.

Richtig Falsch

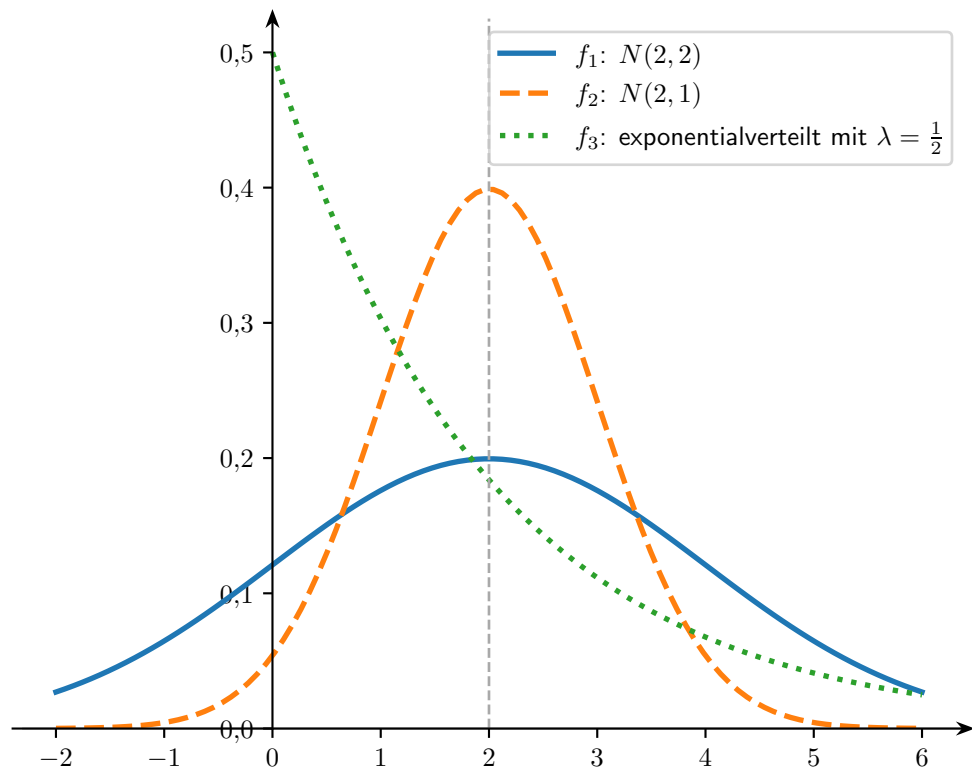
e) Ein Punktschätzer ist ein Intervall, das von der Standardabweichung des Merkmals abhängt.

Richtig Falsch

f) Durch die Aufnahme zweier weiterer Merkmale in die Regressionsgleichung kann das Bestimmtheitsmaß von 0,73 auf 0,76 steigen.

Richtig Falsch

2. (10 Punkte)



Welche der folgenden Aussagen sind richtig?

- a) Alle drei Verteilungen haben denselben Modalwert.
 Richtig Falsch
- b) Verteilungen 1 und 2 sind symmetrisch, Verteilung 3 ist rechtsschief.
 Richtig Falsch
- c) Alle drei Verteilungen haben denselben Erwartungswert.
 Richtig Falsch
- d) Die Verteilungen 1 und 3 haben dieselbe Varianz.
 Richtig Falsch
- e) Die Wahrscheinlichkeit, Werte im Intervall zwischen 2 und 3 zu erhalten, ist bei Verteilung 2 größer als bei Verteilung 1 oder 3.
 Richtig Falsch

3. Der Sommer kommt hoffentlich bald und viele von uns möchten fit in die Sommermonate starten. Dazu wurden 300 Personen gefragt, an wie vielen Tagen sie in der Woche sportliche Aktivitäten ausüben. Das Ergebnis der Umfrage ist in folgender Tabelle zusammengefasst:

Anzahl der Tage	0	1	2	3	4	5	6	7
Absolute Häufigkeit	30	36	54	42	63	33	27	15

- a) i. (5 Punkte) Zeichnen Sie die geeignete Verteilungsfunktion! Bestimmen und interpretieren Sie den Wert $F(3)$!
- ii. (3 Punkte) Berechnen Sie einen geeigneten Lageparameter und begründen Sie ihre Wahl!

Ausführung Beispiel 3a:

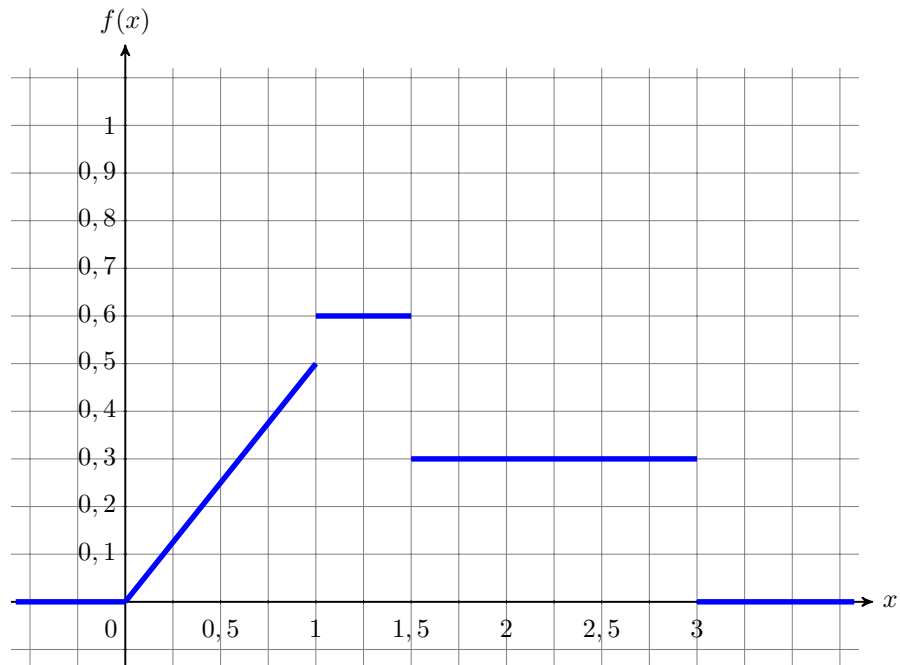
- b) (12 Punkte) Nehmen die sportlichen Aktivitäten mit dem Alter ab? Dazu wurden bei der Umfrage von 6 Personen jeweils das Alter [in Jahren] sowie die Dauer der sportlichen Aktivitäten [in Stunden/Woche] erhoben:

Person	1	2	3	4	5	6
Alter in Jahren	52	41	35	62	22	70
Sportliche Aktivitäten in Stunden/Woche	5	6	5	2	9	3

Mit welcher Kennzahl kann ein Zusammenhang zwischen den beiden Merkmalen am besten dargestellt werden? Berechnen Sie diese und interpretieren Sie Ihr Ergebnis!

Ausführung Beispiel 3b:

4. a) Die Graphik zeigt die Dichtefunktion der Wurfweite bei einem durch einen Spielautomaten simulierten Ballspiel. Mit dieser Dichtefunktion kann die Wahrscheinlichkeit berechnet werden, mit der der Ball in einer bestimmten Entfernung vom Abwurfpunkt auf den Boden fällt.



- i. (4 Punkte) Zeigen Sie, dass diese Funktion eine Dichtefunktion ist!
- ii. (4 Punkte) Bestimmen Sie die Wahrscheinlichkeitsfunktion für die **Auszahlung**, die ein Spieler bei folgendem Spiel erhält:
Fällt der Ball zwischen 0 m und 1 m auf den Boden, bekommt der Spieler nichts. Kommt der Ball zwischen 1 m und 1,5 m auf, erhält der Spieler € 5. Fällt er zwischen 1,5 m und 3 m zu Boden, zahlt der Automat € 7 aus.
- iii. (3 Punkte) Der Einsatz pro Spiel beträgt € 5. Bestimmen Sie den Erwartungswert des **Gewinnes** des Spielers!

- b) Frau A. und Herr B. spielen gleichzeitig an zwei anderen Spielautomaten. Frau A. spielt an einem Gerät, an dem man mit der Wahrscheinlichkeit $\frac{1}{12}$ gewinnt. Herr B. ist weniger risikofreudig und spielt an einem Gerät, an dem man mit der Wahrscheinlichkeit $\frac{1}{9}$ gewinnt, aber eine geringere Auszahlung erhält.
- i. (3 Punkte) Wie groß ist die Wahrscheinlichkeit dafür, dass in der ersten Runde:
 - A. Frau A. gewinnt und Herr B. verliert
 - B. beide verlieren
 - C. mindestens einer gewinnt?
 - ii. (6 Punkte) Wie groß ist die Wahrscheinlichkeit dafür, dass:
 - A. Frau A. genau zwei der nächsten fünf Spiele gewinnt?
 - B. Herr B. mindestens eines der nächsten zehn Spiele gewinnt?

Ausführung Beispiel 4:

5. Ein Landwirtschaftsbetrieb möchte die Erträge zweier verschiedener Tomatensorten – Sorte A und Sorte B – vergleichen. Es werden zwei Stichproben mit jeweils 6 Pflanzen gezogen. Die Erträge (in Kilogramm pro Pflanze) wurden über eine Saison wie folgt erfasst:

Ertrag Sorte A (X)	2	7	5	5	5	6
Ertrag Sorte B (Y)	5	6	8	8	6	3

Nehmen Sie an, dass der Ertrag beider Sorten normalverteilt ist!

- a) i. (6 Punkte) Berechnen Sie ein 95-%-Konfidenzintervall für den durchschnittlichen Ertrag der Sorte A!
ii. (2 Punkte) Interpretieren Sie dieses Intervall im Kontext der Aufgabenstellung!
- b) (4 Punkte) Es soll gezeigt werden, dass der Ertrag der Sorte B signifikant größer ist als der Ertrag der Sorte A.

Formulieren Sie die Nullhypothese (H_0) und die Alternativhypothese (H_1) und entscheiden Sie welcher Test für die Beantwortung dieser Frage durchzuführen wäre! Begründen Sie Ihre Entscheidung verbal, ohne einen Test zu rechnen!

- c) (14 Punkte) Angenommen, die Erträge der Tomaten für mindestens eine der Sorten sind nicht normalverteilt. Welcher Test wäre nun anzuwenden, wenn auf dem 10-%-Niveau gezeigt werden soll, dass der Ertrag der Sorte B signifikant größer ist als der Ertrag der Sorte A? Formulieren Sie die Null- und die Alternativhypothese und führen Sie den Test durch! Interpretieren Sie auch hier das Ergebnis!
- d) (4 Punkte) Sie möchten gerne anhand der gegebenen Stichproben überprüfen, ob die Annahme, dass der Ertrag normalverteilt ist, gerechtfertigt ist. Welcher Test ist hierfür anzuwenden und welche Hypothesen sind aufzustellen?

Ausführung Beispiel 5:

Ausführung Beispiel 5:

6. Mit den meteorologischen Daten der HI-SEAS-Wetterstation soll die Sonneneinstrahlung [Watt pro m²] mittels Regressionsrechnung vorhergesagt werden. Folgende Variablen wurden zur Erklärung verwendet: Temperatur [Grad Fahrenheit]; Luftdruck [mm Hg]; Luftfeuchtigkeit [Prozent]; Windrichtung (Ausprägungen: Nord, Süd, Ost, West); Windgeschwindigkeit [Meilen pro Stunde];

Die Analyse ergab folgendes Ergebnis:

	Koeffizient	Standardfehler	t-Statistik	p-Wert
Schnittpunkt	12.662,132	723,145	17,510	0,000
Temperatur	32,907	0,254	129,559	0,000
Luftdruck	-463,975	23,880	-19,429	0,000
Luftfeuchtigkeit	-0,679	0,049	-13,807	0,000
Windrichtung[Ost]	63,933	3,502	18,256	0,000
Windrichtung[Süd]	-14,716	5,286	-2,784	0,005
Windrichtung[West]	-85,696	2,994	-28,622	0,000
Windgeschwindigkeit	8,979	0,337	26,649	0,000

- (3 Punkte) Um wie viel sinkt/steigt die Sonneneinstrahlung im Schnitt, wenn Ost- statt Westwind herrscht. die Temperatur um 2 Grad Fahrenheit sinkt und alle anderen Variablen gleich bleiben?
- (2 Punkte) Um wie viel sinkt/steigt die Sonneneinstrahlung im Schnitt, wenn Nord- statt Ostwind herrscht und alle anderen Variablen gleich bleiben?
- (3 Punkte) Schätzen Sie mit Hilfe des Modells, wie stark die Sonneneinstrahlung bei einer Temperatur von 50 Grad Fahrenheit, einem Luftdruck von 30,4 mm Hg, einer Luftfeuchtigkeit von 72 % und einer Windgeschwindigkeit von 6 Meilen pro Stunde bei Südwind sein wird!

Ausführung Beispiel 6:

Ausführung Beispiel 6: