

Klausur zur VWA-Statistik

Hinweis: Es sind 4 aus 6 Aufgaben zu bearbeiten. Werden mehr als 4 Aufgaben bearbeitet, werden nur die ersten vier Aufgaben gewertet.

Aufgabe 1:

Betrachten Sie folgenden Datensatz mit Umsätzen in Mio. EUR von 8 Einzelhandelsunternehmen:

Unternehmen	1	2	3	4	5	6	7	8
Umsatz	2.5	10	12.5	1.2	1.0	2.5	8.4	10

1. Zeichnen Sie die empirische Verteilungsfunktion der Umsätze.
2. Berechnen Sie den Median, das untere und das obere Quartil und geben Sie die Spannweite und den Interquartilsabstand an.
3. Berechnen Sie die Varianz der Umsätze.
4. Wie groß ist die relative Häufigkeit der Unternehmen, die mehr als 5 aber weniger als 10.1 Mio. EUR Umsatz machen.

Aufgabe 2:

Für fünf Beschäftigte eines Unternehmens sind folgende Monatsverdienste bekannt:

Beschäftigter i	1	2	3	4	5
Einkommen x_i in EUR	2500	3500	4000	5000	10000

1. Zeichnen Sie die Lorenzkurve.
2. Berechnen Sie das Konzentrationsmaß von Gini, wenn Sie wissen, dass die Summe der absoluten Abweichungen aller Verdienste $\sum_{j=1}^5 \sum_{i=1}^5 |x_i - x_j| = 66000$ beträgt.
3. Jeder Mitarbeiter bekommt dieselbe Gehaltserhöhung von 500 EUR. Wie wirkt sich dies auf die Konzentration aus? Berechnen Sie erneut das Konzentrationsmaß von Gini.
4. Jeder Mitarbeiter bekommt 5 Prozent mehr Gehalt. Wie wirkt sich diese Gehaltserhöhung auf die Konzentration aus?

Aufgabe 3:

Für 8 Unternehmen, die alle das gleiche Produkt herstellen, liegen Angaben über ihre jährliche Produktion (in Tsd. Stück) und ihre Durchschnittskosten (in EUR) vor. Untersuchen Sie, ob es einen Zusammenhang zwischen der Höhe der Produktion und den Durchschnittskosten gibt.

Unternehmen i	1	2	3	4	5	6	7	8
Produktion in Tsd. Stück x_i	1000	1500	800	2000	700	10000	8000	2000
Durchschnittskosten in EUR y_i	0.12	0.15	0.25	0.11	0.30	0.08	0.10	0.12

Sie können benutzen, dass $\sum_{i=1}^8 x_i = 26000$, $\sum_{i=1}^8 x_i^2 = 176380000$, Mittelwert bzw. Varianz der Stückkosten 0.15375 bzw. 0.005398438 und $\sum_{i=1}^8 x_i y_i = 2815$ betragen.

1. Tragen Sie die Wertepaare in ein Streuungsdigramm (Scatterplot) ein. Vermuten Sie einen positiven oder einen negativen Zusammenhang zwischen Produktionsmenge und Stückkosten? (Begründung)
2. Berechnen Sie den Korrelationskoeffizienten und interpretieren Sie dessen Wert.
3. Berechnen Sie Achsenabschnitt \hat{a}_0 und Steigung \hat{a}_1 der Regressionsgeraden

$$\text{Stückkosten} = \hat{a}_0 + \hat{a}_1 \text{Produktion.}$$

4. Das Bestimmtheitsmaß dieser Regression beträgt 0.3524. Interpretieren Sie diesen Wert.

Aufgabe 4:

Ein Würfel mit vier Seiten ist eine Pyramide. Zusätzlich wird eine Münze betrachtet, deren Rand so breit ist, dass die Münze auch auf dem Rand liegen bleiben kann. Die Seiten der Pyramide sind von 1 bis 4 und die Seiten der Münze von 1 bis 3 durchnummeriert. Der Rand der Münze trage die Zahl 3. Er soll mit einer Wahrscheinlichkeit von 20% beim Münzwurf auftreten. Als Zufallsexperiment wird betrachtet, dass zuerst die Münze und dann die Pyramide geworfen wird. Die Augenzahlen des Münz- X und des Pyramidenwurfes Y seien stochastisch unabhängig.

1. Geben Sie den Stichprobenraum (= Menge aller Elementarereignisse an).
2. Berechnen Sie die folgenden Wahrscheinlichkeiten:
 - $P(X = 1)$
 - $P(X = 3)$
 - $P(X = 1, Y = 2)$
 - $P(X = 3, Y = 2)$
3. Berechnen Sie die Wahrscheinlichkeit, dass die Augensumme $X + Y$ gerade ist.
4. Berechnen Sie die Wahrscheinlichkeit, dass das Produkt der Augenzahlen $X \cdot Y$ kleiner oder gleich 3 ist.
5. Berechnen Sie die Wahrscheinlichkeit, dass die Münze auf dem Rand stehen bleibt unter der Bedingung, dass die Augensumme gerade ist.
6. Sind das Ereignis, die Münze bleibt auf dem Rand stehen und das Ereignis, die Augensumme $X + Y$ ist gerade, stochastisch unabhängig? (Begründung)

Aufgabe 5:

Eine Publikation soll auf Plagiate überprüft werden. Dazu sollen zufällig n Seiten ausgewählt und auf Plagiate durchsucht werden. Die Anzahl der Plagiate pro Seite sei ungefähr normalverteilt mit unbekanntem Mittelwert μ und Varianz $\sigma^2 = 0.04$.

1. Bestimmen Sie den Stichprobenumfang n so, dass die mittlere Anzahl von Plagiaten in der Stichprobe vom unbekanntem Mittelwert mit einer Wahrscheinlichkeit von mindestens 95% um weniger als 0.1 abweicht.
2. Welche der folgenden Aussagen sind richtig? (Begründung)
 - (a) Der benötigte Stichprobenumfang wird kleiner, wenn die Sicherheitswahrscheinlichkeit steigt, mit der der Stichprobenmittelwert vom wahren Mittelwert um weniger als 0.1 abweicht.
 - (b) Der benötigte Stichprobenumfang wird größer, wenn die Differenz größer gewählt wird, um die der Stichprobenmittelwert vom wahren Mittelwert maximal abweichen darf.
 - (c) Die Varianz der Grundgesamtheit σ^2 hat keinen Einfluss auf den benötigten Stichprobenumfang.

Aufgabe 6:

In Haslach betreibt die Gesellschaft für Konsumforschung (GfK) einen sog. Testmarkt. D.h. es werden dort gezielt bei einem Teil der Probanden Werbespots in die Fernsehsendungen eingespielt und anschließend beobachtet, ob die Spots das Kaufverhalten bezüglich des beworbenen Produktes beeinflussen.

Es liegen 101 Käufe von Kunden vor, die den Werbespot gesehen haben und 81 Käufe von Kunden vor, die den Werbespot nicht gesehen haben.

Für diese unverbundenen Stichproben soll mit einer Irrtumswahrscheinlichkeit von 5% getestet werden, ob der mittlere Umsatz μ_1 der Kunden mit Werbespot größer ist als der mittlere Umsatz μ_2 der Kunden ohne Werbespot. Es kann davon ausgegangen werden, dass der Umsatz ungefähr normalverteilt ist.

Es sei Ihnen bekannt:

1. Stichprobenmittelwert des Umsatzes mit Werbespot: 10 EUR
2. Stichprobenvarianz des Umsatzes mit Werbespot: 2 EUR²,
3. Stichprobenmittelwert des Umsatzes ohne Werbespot: 8 EUR
4. Stichprobenvarianz des Umsatzes ohne Werbespot: 3 EUR².

Stellen Sie das zu testende Hypothesenpaar auf, geben Sie die Entscheidungsregel allgemein an und leiten Sie die Testentscheidung für eine Irrtumswahrscheinlichkeit von 5% her.

Hinweis: Sie können bei den vorliegenden Stichprobenumfängen die Quantile der Standardnormalverteilung benutzen.

Quantile der Standardnormalverteilung

<i>P</i>	.0	.1	.2	.3	.4	.5	.6	.7	.8	.9
75	.6745	.6776	.6808	.6840	.6871	.6903	.6935	.6967	.6999	.7031
76	.7063	.7095	.7127	.7160	.7192	.7225	.7257	.7290	.7323	.7356
77	.7388	.7421	.7454	.7488	.7521	.7554	.7588	.7621	.7655	.7688
78	.7722	.7756	.7790	.7824	.7858	.7892	.7926	.7961	.7995	.8030
79	.8064	.8099	.8134	.8169	.8204	.8239	.8274	.8310	.8345	.8381
80	.8416	.8452	.8488	.8524	.8560	.8596	.8632	.8669	.8705	.8742
81	.8779	.8816	.8853	.8890	.8927	.8965	.9002	.9040	.9078	.9116
82	.9154	.9192	.9230	.9269	.9307	.9346	.9385	.9424	.9463	.9502
83	.9542	.9581	.9621	.9661	.9701	.9741	.9781	.9822	.9863	.9904
84	.9945	.9986	1.0027	1.0069	1.0110	1.0152	1.0194	1.0237	1.0279	1.0322
85	1.0364	1.0407	1.0450	1.0494	1.0537	1.0581	1.0625	1.0669	1.0714	1.0758
86	1.0803	1.0848	1.0893	1.0939	1.0985	1.1031	1.1077	1.1123	1.1170	1.1217
87	1.1264	1.1311	1.1359	1.1407	1.1455	1.1503	1.1552	1.1601	1.1650	1.1700
88	1.1750	1.1800	1.1850	1.1901	1.1952	1.2004	1.2055	1.2107	1.2160	1.2212
89	1.2265	1.2319	1.2372	1.2426	1.2481	1.2536	1.2591	1.2646	1.2702	1.2759
90	1.2815	1.2873	1.2930	1.2988	1.3047	1.3106	1.3165	1.3225	1.3285	1.3346
91	1.3408	1.3469	1.3532	1.3595	1.3658	1.3722	1.3787	1.3852	1.3917	1.3984
92	1.4051	1.4118	1.4187	1.4255	1.4325	1.4395	1.4466	1.4538	1.4611	1.4684
93	1.4758	1.4833	1.4908	1.4985	1.5063	1.5141	1.5220	1.5301	1.5382	1.5464
94	1.5548	1.5632	1.5718	1.5805	1.5893	1.5982	1.6072	1.6164	1.6258	1.6352
95	1.6448	1.6546	1.6646	1.6747	1.6849	1.6954	1.7060	1.7169	1.7279	1.7392
96	1.7507	1.7624	1.7744	1.7866	1.7991	1.8119	1.8250	1.8384	1.8522	1.8663
97	1.8808	1.8957	1.9110	1.9268	1.9431	1.9600	1.9774	1.9954	2.0141	2.0335
98	2.0537	2.0748	2.0969	2.1201	2.1444	2.1701	2.1973	2.2262	2.2571	2.2904
99	2.3263	2.3656	2.4089	2.4572	2.5121	2.5758	2.6520	2.7477	2.8781	3.0901