Klausur: Mathematik und Statistik

Lehrveranstaltung: Statistik

Fakultät für Wirtschaft

Studiengang: BWL-Öffentliche Wirtschaft

Datum: 27.06.2022

Studierende Matrikelnummer:

Dozent/in: Jürgen Meisel

Kurs: WOW21A/B **Semester:** 2.

Hilfsmittel: ***Wiss. TR (nicht programmierbar) / Formelsammlung***

Bearbeitungszeit: 60 Minuten

Bewertung: **Maximale Punktzahl:** 60 Punkte

Erreichte Punktzahl:

Datum, Unterschrift

……………………………………..

Anmerkungen:

Von 7 gestellten Aufgaben müssen 5 ausgewählt und bearbeitet werden.

Auf-gabennr.:	Thema / Bereich	maximale Punkte	erreichte Punkte	Bemerkungen
1	Mittelwerte & Streumaße (diskret)	12		
2	Mittelwerte & Streumaße (klassiert)	12		
3	Konzentration (Ginikoeff. & Lorenzkurve)	12		
4	Regression & Korrelation	12		
5	Preisindizes und Inflationsrate	12		
6	Stochastik: Verteilung (diskret)	12		
7	Stochastik: Verteilung (stetig)	12		
		5 aus 7		
Summe		**60**		

Klausur: Statistik – 27.06.2022

Jürgen Meisel

Hilfsmittel: **Wiss. nicht progr. Taschenrechner + Formelsammlung**
Bearbeitungszeit: **60 Minuten**

Aufgabe 1: **Mittelwerte und Streumaße (diskret)**

Teil 1:
Im letzten Winter haben Sie intensiv Orangen gekauft – insgesamt 150 Stück – und diese jeweils gewogen 😉
Die nachfolgende Tabelle zeigt die relative Häufigkeitsverteilung des Gewichts:

Gewicht in g	234	243	249	250	253	257
relative Häufigkeit	0.1	0.1	0.1	0.2	0.3	0.2

a) Geben Sie die absoluten Häufigkeiten für das jeweilige Gewicht an.
b) Bestimmen Sie den arithmetischen Mittelwert, den Modalwert und den Median.
c) Wie hoch sind die Standardabweichung, das untere Quartil und das obere Quartil.

Lösung:

Gewicht	234	243	249	250	253	257
Rel. H‘keit	0,1	0,1	0,1	0,2	0,3	0,2
Abs. H’keit	15	15	15	30	45	30

Mittelwert:

$$\bar{x} = \sum_{i=1}^{k}(x_i \cdot p_i)$$

$$\bar{x} = 234\cdot 0{,}1 + 243\cdot 0{,}1 + 249\cdot 0{,}1 + 250\cdot 0{,}2 + 253\cdot 0{,}3 + 257\cdot 0{,}2 = 249{,}9$$

Median: $$\overline{x_M} = \frac{1}{2}\left(x_{\frac{n}{2}} + x_{\frac{n}{2}+1}\right) \overset{n=150}{=} \frac{1}{2}(x_{75}+x_{76}) = \frac{1}{2}(250+253) = 251{,}5$$

Modalwert: Häufigster Wert: 253

Standardabweichung:

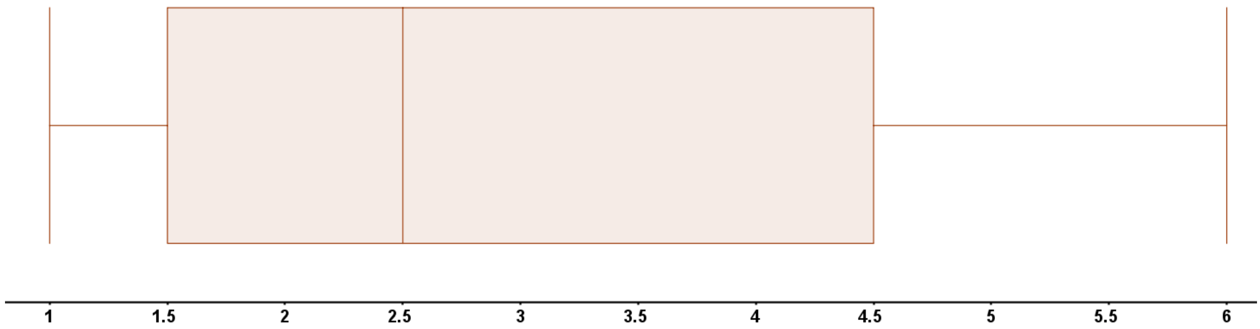
$$V(X) = \left[\sum_{i=1}^{n} x_i^2 \cdot p_i\right] - \bar{x}^2 \rightarrow V(X) = 43{,}09 \xrightarrow{\sqrt{\ }} S(X) = 6{,}564$$

Quartil 1: $$\overline{x_p} \underset{150\cdot 0{,}25=37{,}5}{\overset{n\cdot p}{=}} x_{[n\cdot p]+1} \xrightarrow{\text{Quartil 1}} \overline{x_{0,25}} = x_{[37{,}5]+1} = x_{38} = 249$$

Quartil 3: $$\overline{x_p} \underset{150\cdot 0{,}75=112{,}5}{\overset{n\cdot p}{=}} x_{[n\cdot p]+1} \xrightarrow{\text{Quartil 3}} \overline{x_{0,75}} = x_{[112{,}5]+1} = x_{113} = 253$$

Teil 2:

Gegeben sei ein Boxplot auf der Basis einer Kursarbeit in Mathematik mit insgesamt 32 Beteiligten.



Ermitteln Sie folgende Noten aus der (fehlenden) tabellierten Noten-Verteilung mit kurzer Begründung:

x_8 **=>** $Ganzzahlig: \quad 32 \cdot 0,25 = 8 \quad \rightarrow \quad q_1 = \frac{1}{2}(x_8 + x_9) = 1,5 \quad \rightarrow \quad x_8 = 1$

x_{16} **=>** $n = 32: \quad \rightarrow \quad \overline{x_M} = \frac{1}{2}(x_{16} + x_{17}) = 2,5 \quad \rightarrow \quad x_{16} = 2$

x_{25} **=>** $Ganzzahlig: \quad 32 \cdot 0,75 = 24 \quad \rightarrow \quad q_3 = \frac{1}{2}(x_{24} + x_{25}) = 4,5 \quad \rightarrow \quad x_{25} = 5 \; oder \; 6$

Aufgabe 2: **Mittelwerte und Streumaße (klassiert)**

In der Personalabteilung der Firma „Rasch und Ruh – Morgens geschlossen, mittags zu“ ist zum 31.05.2022 die Altersstruktur der Mitarbeiter zu ermitteln gewesen.

Häufigkeitstabelle der Mitarbeiter (nach Altersstufen):

Alter [Jahre]	absolute Häufigkeit	relative Häufigkeit	Klassen-mitte	Klassen-breite	Häufigkeits-dichte	kum. rel. Häufigkeit
[16 ; 20[	5					
[20 ; 30[	18					
[30 ; 40[	13					
[40 ; 50[	24					
[50 ; 60[	12					
[60 ; 66]	8					
Summe	80					

a) Vervollständigen Sie die Tabelle.
b) Bestimmen Sie den arithmetischen Mittelwert, die modale Klasse und den Modalwert.
c) Berechnen Sie den Median, das untere Quartil und das obere Quartil.

Lösung:

Häufigkeitstabelle der Mitarbeiter (nach Altersstufen):

Alter [Jahre]	absolute Häufigkeit	relative Häufigkeit	Klassen-mitte	Klassen-breite	Häufigkeits-dichte	kum. rel. Häufigkeit
[16 ; 20[	5	*0,0625*	*18*	*4*	*1,25*	*0,0625*
[20 ; 30[	18	*0,225*	*25*	*10*	*1,8*	*0,2875*
[30 ; 40[	13	*0,1625*	*35*	*10*	*1,3*	*0,45*
[40 ; 50[	24	*0,3*	*45*	*10*	*2,4*	*0,75*
[50 ; 60[	12	*0,15*	*55*	*10*	*1,2*	*0,9*
[60 ; 66]	8	*0,1*	*63*	*6*	*1,33*	*1*
Summe	80	*1*		*50*		

$$\bar{x} = 18\cdot\frac{5}{80}+25\cdot\frac{18}{80}+35\cdot\frac{13}{80}+45\cdot\frac{24}{80}+55\cdot\frac{12}{80}+63\cdot\frac{8}{80} = 40,4875 \approx 40,5\left[Jahre\right]$$

$$Modale\ Klasse: \ \left[40\,;50\right] wegen\ HD_{\max} = 2,4 \ \rightarrow \ Modalwert: \overline{x_{Mod}} = 45$$

$$\overline{x_{Median}} = 40+\frac{10\left(0,5-0,45\right)}{0,3} = 41,67 \qquad \overline{x_{quartile\,1}} = 20+\frac{10\left(0,25-0,0625\right)}{0,225} = 28,33$$

$$\overline{x_{quartile\,3}} = 40+\frac{10\left(0,75-0,45\right)}{0,3} = 50$$

Aufgabe 3: **Gini-Koeffizient & Lorenzkurve**

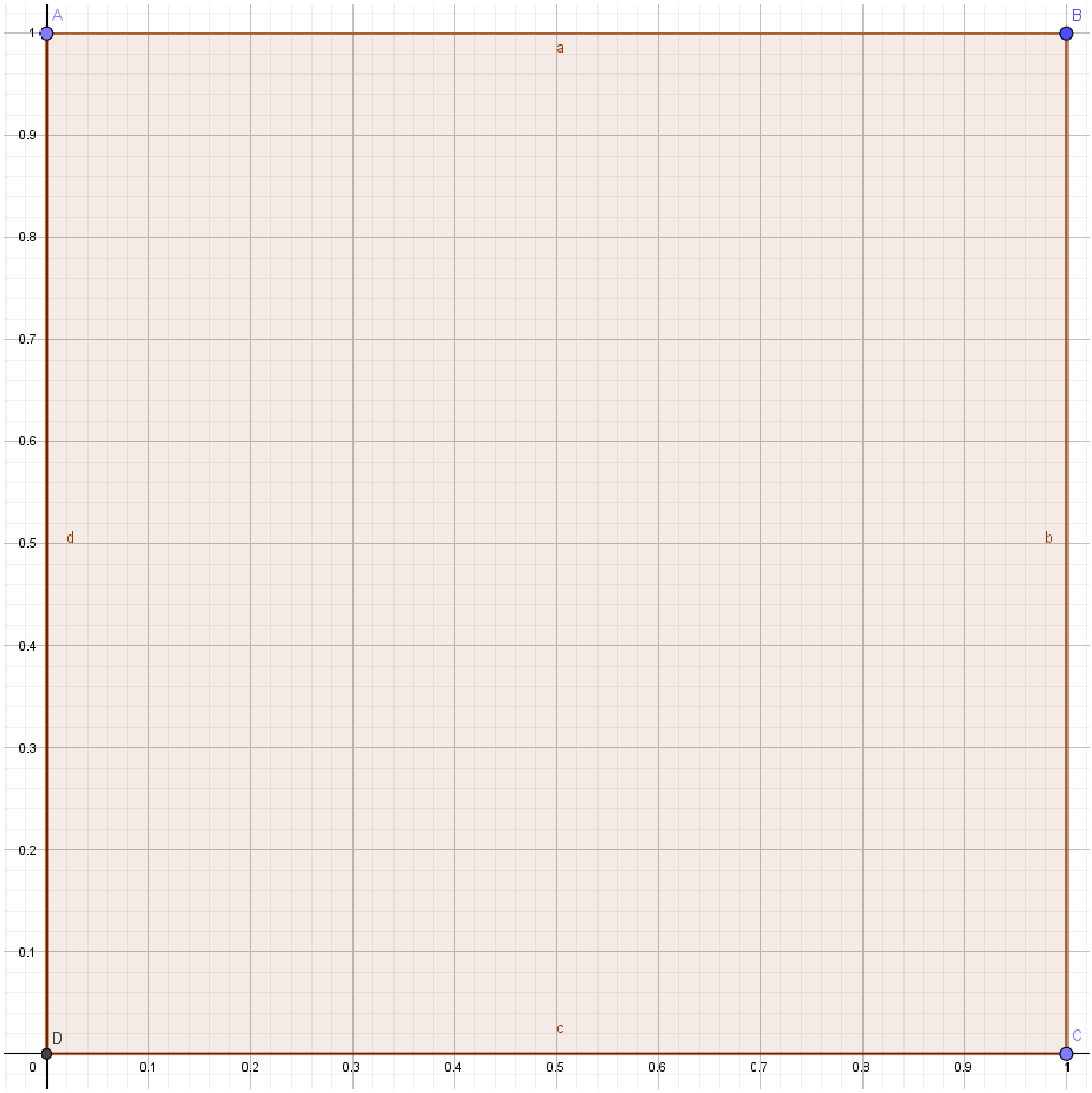
Ein Markt wird von **fünf Unternehmen** beliefert. **3 Unternehmen besitzen jeweils 10 % Lieferanteil**, **der vierte Lieferant hat einen Anteil von 20 %**, der **fünfte vervollständigt das Kontingent**.

a) Zeichnen Sie die dazugehörige Lorenzkurve.

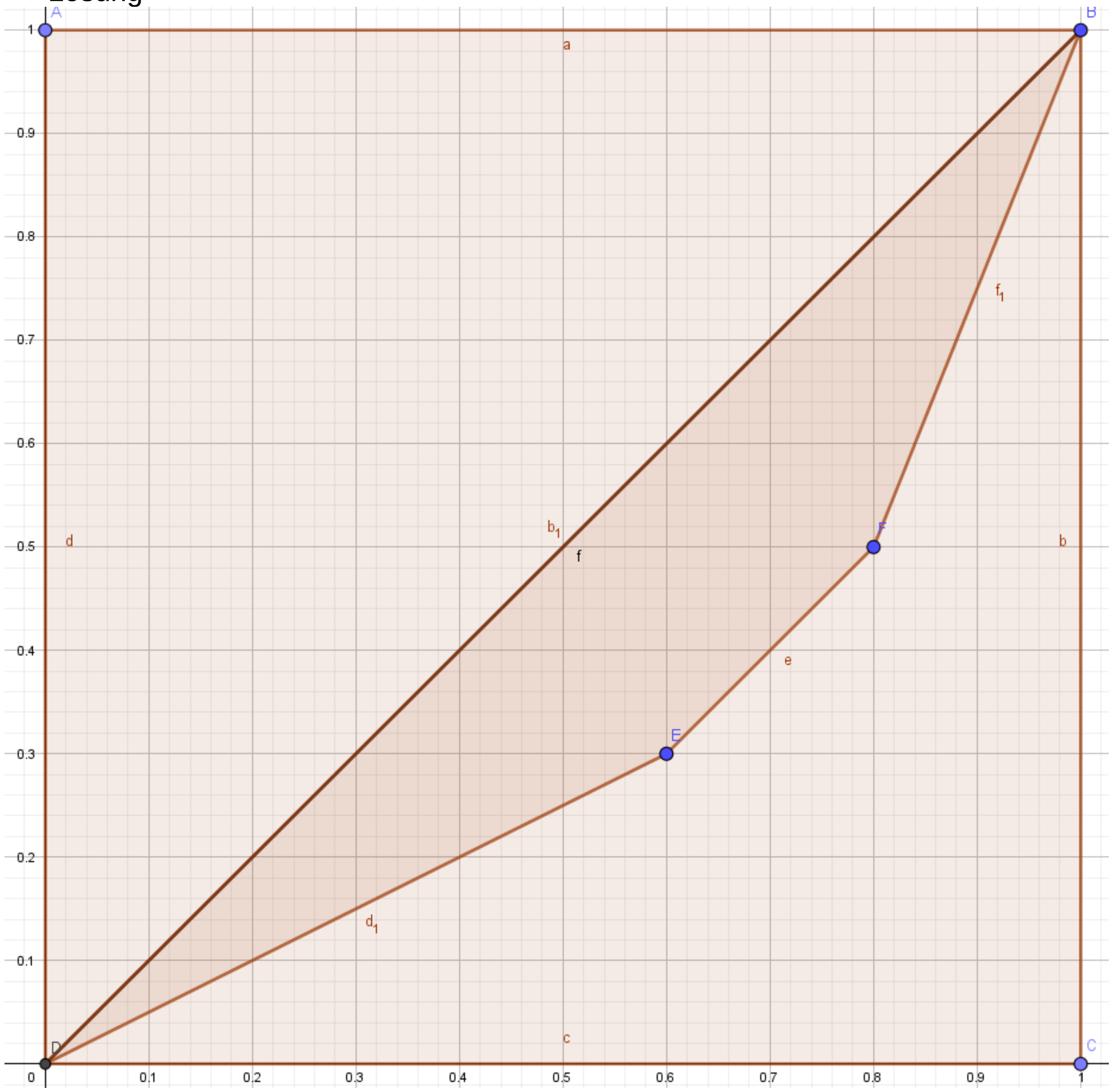
b) Berechnen Sie den Gini-Koeffizienten und den normierten Gini-Koeffizienten.

c) Warum ist die Differenz zwischen den beiden Gini-Koeffizienten relativ groß?

Anlage zu a)



Lösung



Fläche unterhalb der Lorenzkurve [FuL]: FuL = 0,09+0,08+0,15 = 0,32

⇨ Konzentrationsfläche [KF]: KF = 0,18

⇨ Gini-Koeffizient [GK]: GK = KF/0,5 = 0,36

⇨ Normierter Gini-Koeffizient: $GK_{(norm.)}$ = KF * 2n/(n-1) = 0,18 * 10/4 = 0,45

⇨ Problem des großen Unterschieds: geringe Grundmenge n = 5

Aufgabe 4: Regression & Korrelation

Bei einem Experiment wurde bei 15 Studenten ein IQ-Test durchgeführt und der jeweilige Intelligenzquotient (X) berechnet. Zudem wurden die Teilnehmer gebeten, auf einer Skala von 1 (sehr schlecht) bis 9 (sehr gut) die Qualität ihrer persönlichen Leistung (Y) einzuschätzen, ohne das Ergebnis der IQ-Berechnung zuvor erfahren zu haben.

In der Tabelle sind die entsprechenden Werte dargestellt:

i	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
X	70	98	85	82	95	75	93	65	90	77	107	104	124	113	117
Y	3	8	5	4	7	1	6	2	9	4	9	7	9	5	6

Sie möchten nun folgende Hypothese prüfen:

„Die Probanten sind sehr gut in der Lage, ihre Leistung im IQ-Test einzuschätzen."

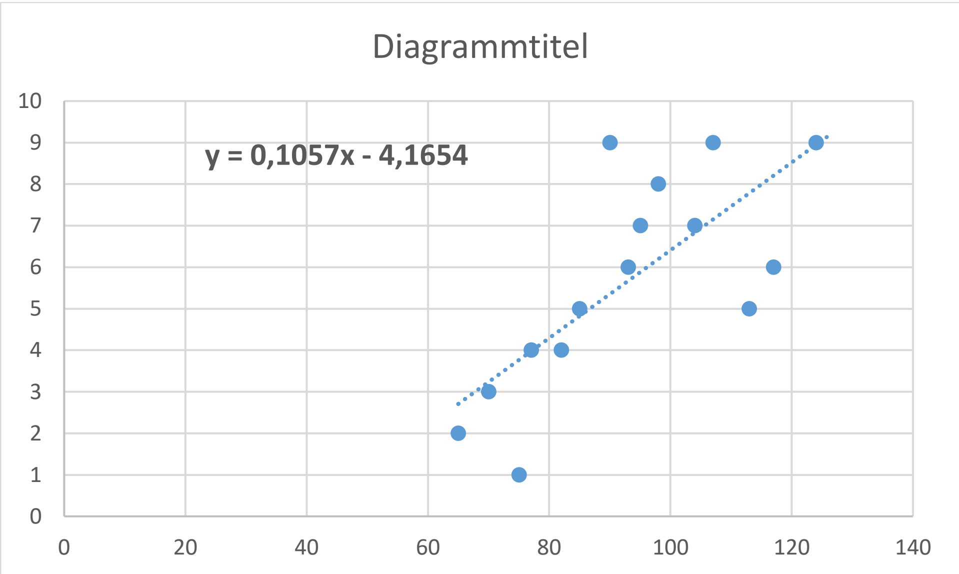
a) Bestimmen Sie die lineare Regressionsfunktion aus IQ und Selbsteinschätzung.
b) Ermitteln Sie die ***Korrelation nach Pearson* und *die Kovarianz*** zwischen IQ und Qualität der Leistungseinschätzung.
c) Beurteilen Sie die Hypothese entsprechend des Korrelationswertes?

Lösung:

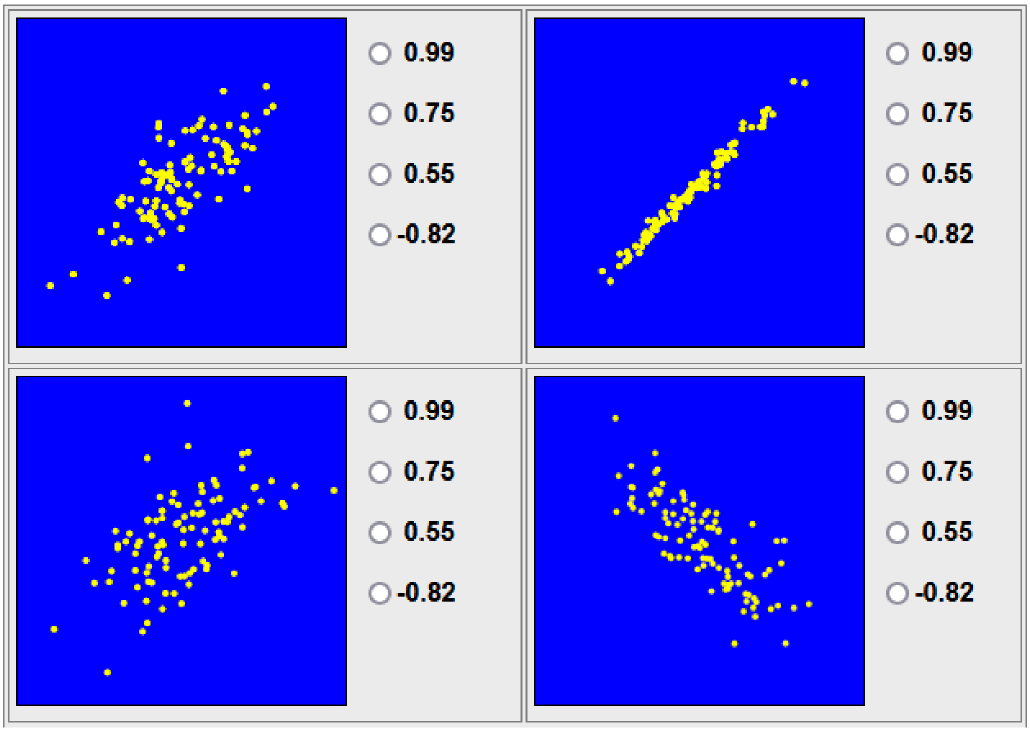
Achsenabschnitt:		**-4,16536995**
Steigung:		**0,10572082**
Korrelation:		**0,7312849**
Kovarianz		**30,8**
Standardabw. X:		**17,0684895**
Standardabw. Y:		**2,4675674**

$$Cov(X,Y) \overset{Def.}{=} \frac{1}{n}\sum_{i=1}^{n}(x_i-\bar{x})(y_i-\bar{y})$$

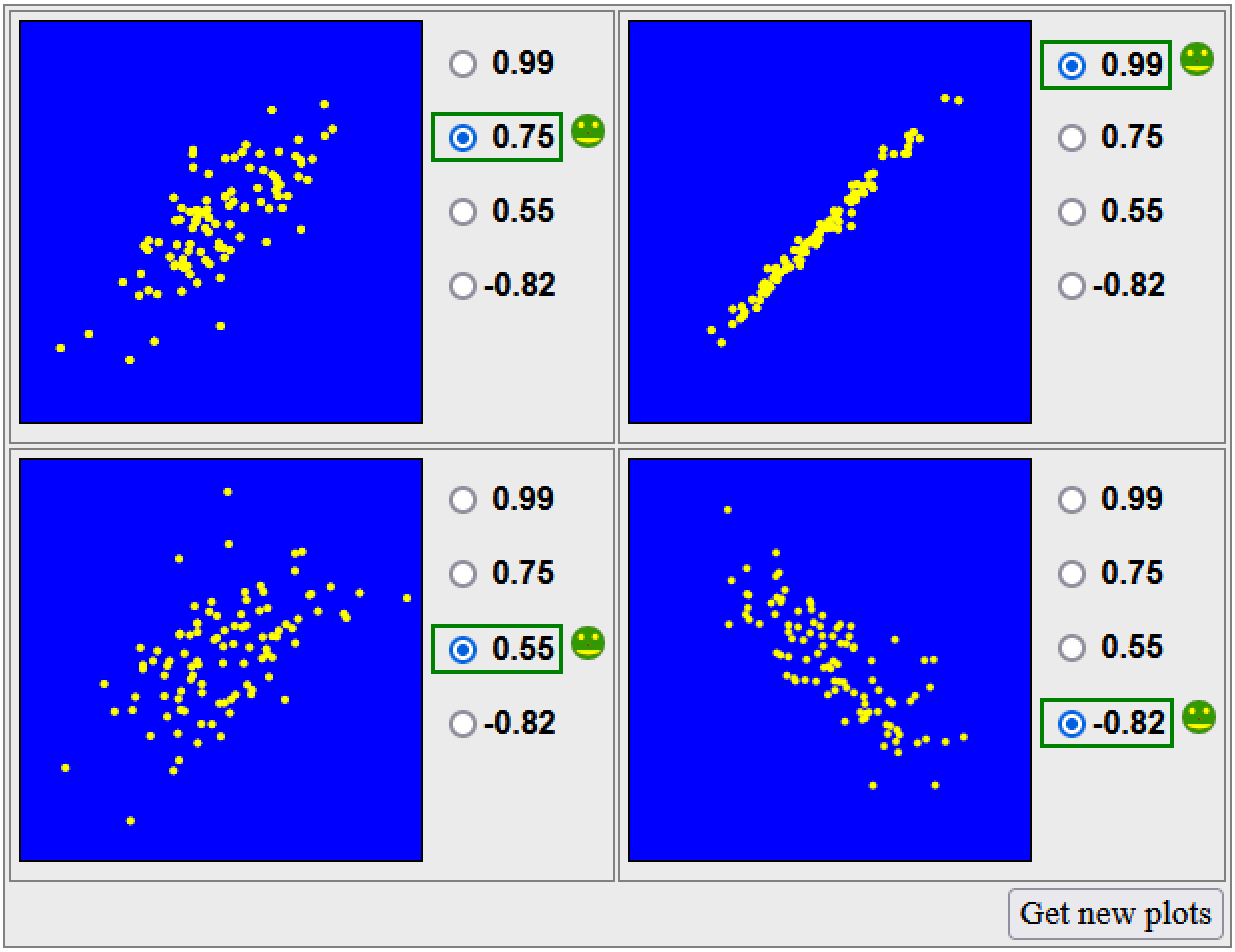
$$Cov(X,Y) \overset{Berechnung}{=} \frac{1}{n}\sum_{i=1}^{n}(x_i\cdot y_i) - \bar{x}\cdot\bar{y}$$



d) Ordnen Sie die „Punktwolken“ den korrekten Werten der Korrelationskoeffizienten zu.



Lösung:



Round 1 result: 4 correct. Total so far: 4 out of 4 = 100%.

Guessing Correlations\. Sie gelangen zu dem Applet wie folgt: http://istics.net/Correlations/

Aufgabe 5: **Warenkorbmethode und Preisindexberechnung**

Ein Unternehmen hat eine Preis-Mengen-Übersicht für die bezogenen Güter A, B und C angefertigt.

Gut	Preise		Mengen	
	2015	2021	2015	2021
A	10	15	60	50
B	25	20	40	70
C	30	40	80	60

a) Ermitteln Sie hierzu die Preisindices nach Laspeyres und Paasche.

b) Berechnen Sie den Preisindex nach Fisher.

c) Wie hoch ist die jährliche Inflationsrate auf der Grundlage der Daten nach Laspeyres?

Lösung:

$$Laspeyres: \quad L_P = \frac{\sum p_{21i} \cdot q_{15i}}{\sum p_{15i} \cdot q_{15i}}$$

$$\rightarrow \frac{\text{Ausgaben des Berichtsjahres mit Mengen des Basisjahres (Menge }_{\text{Periode I}} \text{ * Preis }_{\text{Periode II}})}{\text{Ausgaben/Umsatz des Basisjahres (Menge }_{\text{Periode I}} \text{ * Preis }_{\text{Periode I}})}$$

$$L_P = \frac{15 \cdot 60 + 20 \cdot 40 + 40 \cdot 80}{10 \cdot 60 + 25 \cdot 40 + 30 \cdot 80} = \frac{4900}{4000} = 1,225$$

$$Paasche: \quad P_P = \frac{\sum p_{21i} \cdot q_{21i}}{\sum p_{15i} \cdot q_{21i}}$$

$$\rightarrow \frac{\text{Ausgaben/Umsatz des Berichtsjahres (Menge }_{\text{Periode II}} \text{ * Preis }_{\text{Periode II}})}{\text{Ausgaben des Basisjahres mit Mengen des Berichtsjahres (Menge }_{\text{Periode II}} \text{ * Preis }_{\text{Periode I}})}$$

$$P_P = \frac{15 \cdot 50 + 20 \cdot 70 + 40 \cdot 60}{10 \cdot 50 + 25 \cdot 70 + 30 \cdot 60} = \frac{4550}{4050} = 1,1234$$

$$F_P = \sqrt{L_P \cdot P_P} \quad \rightarrow \quad \sqrt{1,225 \cdot 1,1234} = 1,1731$$

$$\text{Inflationsrate:} \quad \sqrt[6]{1,225} = 1,034402 \rightarrow i_{eff} = 1,034402 - 1 = 0,034402 \xrightarrow{\cdot 100} 3,4402\%$$

d) Für einen aus 400 Gütern bestehenden Warenkorb wurden für die Jahre 2018, 2019 und 2020 folgende Umsatzsummen berechnet:

$$\sum_{i=1}^{400} q_{18;i}\, p_{18;i} = 870 \qquad \sum_{i=1}^{400} q_{18;i}\, p_{19;i} = 877 \qquad \sum_{i=1}^{400} q_{18;i}\, p_{20;i} = 898$$

$$\sum_{i=1}^{400} q_{19;i}\, p_{18;i} = 873 \qquad \sum_{i=1}^{400} q_{19;i}\, p_{19;i} = 879 \qquad \sum_{i=1}^{400} q_{19;i}\, p_{20;i} = 902$$

$$\sum_{i=1}^{400} q_{20;i}\, p_{18;i} = 878 \qquad \sum_{i=1}^{400} q_{20;i}\, p_{19;i} = 895 \qquad \sum_{i=1}^{400} q_{20;i}\, p_{20;i} = 905$$

Berechnen Sie hieraus die Preisindices nach Laspeyres und Paasche für 2020 zum Basisjahr 2018.

Lösung

$$Laspeyres: \quad L_P = \frac{\sum p_{20i} \cdot q_{18i}}{\sum p_{18i} \cdot q_{18i}} = \frac{898}{870} = 1{,}032184$$

$$\rightarrow \frac{\text{Ausgaben des Berichtsjahres mit Mengen des Basisjahres (Menge}_{\text{ Periode I}} \text{ * Preis}_{\text{ Periode II}})}{\text{Ausgaben/Umsatz des Basisjahres (Menge}_{\text{ Periode I}} \text{ * Preis}_{\text{ Periode I}})}$$

$$Paasche: \quad P_P = \frac{\sum p_{20i} \cdot q_{20i}}{\sum p_{18i} \cdot q_{20i}} = \frac{905}{878} = 1{,}030752$$

$$\rightarrow \frac{\text{Ausgaben/Umsatz des Berichtsjahres (Menge}_{\text{ Periode II}} \text{ * Preis}_{\text{ Periode II}})}{\text{Ausgaben des Basisjahres mit Mengen des Berichtsjahres (Menge}_{\text{ Periode II}} \text{ * Preis}_{\text{ Periode I}})}$$

Aufgabe 6: **Stochastik – Binomialverteilung - Orangenkauf**

Im Großhandel weiß man, dass die Wahrscheinlichkeit für eine Orange in schlechtem Zustand **P(S) = 0,1 bzw. 10 %** beträgt.
Ein Einzelhändler führt eine Testserie durch und prüft **80 Orangen**.

a) Wie hoch ist der **Erwartungswert** für Orangen in gutem Zustand?

b) Mit welcher Wahrscheinlichkeit beinhaltet die Testserie
   (i) **Höchstens 3 schlechte** Orangen?
   (ii) **Mindestens 70 einwandfreie** Orangen?

Lösung:

$$\mu_{BV}(X) = n \cdot p \rightarrow \mu = 80 \cdot 0{,}9 = 72$$

$$B_{80;0,1}(X \leq 3) = \sum_{X=0}^{3} \binom{80}{X} 0{,}1^{X} \cdot 0{,}9^{80-X} = 0{,}0353$$

$$B_{80;0,9}(Y \geq 70) = \sum_{Y=70}^{80} \binom{80}{Y} 0{,}9^{Y} \cdot 0{,}1^{80-Y} = 0{,}8266$$

Die Orangen sind in Kisten zu **80 Stück** verpackt.
Eine **Kiste kann eingelagert** werden, wenn ***höchstens 3 Orangen schlecht*** sind.

c) Lieferant Hummel bestellt **10 Kisten** und möchte nun wissen, wie hoch die Wahrscheinlichkeit ist, dass er **mind. 8 davon einlagern** kann.

Lösung:

$$P(\text{"}einlagern\text{"}) = B_{80;0,1}(X \leq 3) = \sum_{X=0}^{3} \binom{80}{X} 0{,}1^{X} \cdot 0{,}9^{80-X} = 0{,}0353$$

$$B_{10;0,9647}(X \geq 8) = \sum_{X=8}^{10} \binom{10}{X} 0{,}9647^{X} \cdot 0{,}0353^{10-X} = 0{,}9956$$

Aufgabe 7: **Stochastik –Normalverteilung - Limettenkauf**

Es wird eine Bestellung von 10.000 Limetten durchgeführt. Der Lieferant gibt die Zusage, dass die Limetten zu 98 % einwandfrei sind.

a) Ermitteln Sie das 95%-Intervall innerhalb dessen Limetten guter Qualität zu erwarten sind.

$$\mu(X) = n \cdot p \rightarrow \mu = 10.000 \cdot 0{,}98 = 9.800$$

$$\sigma(X) = \sqrt{n \cdot p \cdot (1-p)} \rightarrow \sigma(X) = \sqrt{10.000 \cdot 0{,}98 \cdot 0{,}02} = 14$$

$$z = \frac{k-\mu}{\sigma} \rightarrow 1{,}96 = \frac{k-9800}{14} \rightarrow k_2 = 9.827{,}44 \xrightarrow{9800-27{,}44} k_1 = 9.772{,}56$$

$$\text{Intervall: } [9.772{,}56\,;\,9.827{,}44]$$

b) Wie hoch ist die Wahrscheinlichkeit höchstens 120 Limetten mit schlechter Qualität zu erhalten?

$$\mu(X) = n \cdot p \rightarrow \mu = 10.000 \cdot 0{,}02 = 200$$

$$\sigma(X) = \sqrt{n \cdot p \cdot (1-p)} \rightarrow \sigma(X) = \sqrt{10.000 \cdot 0{,}02 \cdot 0{,}98} = 14$$

$$P(X \leq 120) = \Phi_{\mu;\sigma}\left(\frac{120-200}{14}\right) = \Phi_{\mu;\sigma}(-5{,}714) = 1-\Phi_{\mu;\sigma}(5{,}714) \approx 0$$

$$z = \frac{120-200}{14} = -5{,}714$$

c) Wie viele Limetten schlechter Qualität dürfen höchstens in der Lieferung enthalten sein, wenn die verbesserte Zusage des Lieferanten gelten soll:
Höchstens 1% der Lieferung ist nicht verwertbar!!!

$$\mu(X) = n \cdot p \rightarrow \mu = 10.000 \cdot 0{,}01 = 100$$

$$\sigma(X) = \sqrt{n \cdot p \cdot (1-p)} \rightarrow \sigma(X) = \sqrt{10.000 \cdot 0{,}01 \cdot 0{,}99} = 9{,}95$$

$$\rightarrow \Phi_{\mu;\sigma}\left(\frac{k-100}{9{,}95}\right) \leq 0{,}01 \rightarrow -2{,}33 = \frac{k-100}{9{,}95} \rightarrow k = 76{,}8165$$