Jürgen Meisel
FH Ludwigshafen

Klausur: Statistik
am Montag, den 30. Januar 2006

Zugelassene Hilfsmittel: Formelsammlung Statistik und nicht progr. Taschenrechner
Bearbeitungszeit: **120 Minuten**

Statistik

1.) **Mittelwerte und Streumaße: Dauer der Vereinszugehörigkeit**

Anlässlich des 30-jährigen Jubiläums des Hühnervereins Eihausen wurde die Zugehörigkeit der Mitglieder festgestellt.
Es ergab sich folgende Verteilung:

Jahre	[0;1]	]1;3]	]3;7]	]7;10]	]10;15]	]15;20]	]20;30]
Anzahl	5	8	10	8	9	12	8

a) Bestimmen Sie die durchschnittliche Dauer der Vereinszugehörigkeit.

b) Errechnen Sie die Varianz und die Standardabweichung.

Lösung:

a) $$\bar{x} = \frac{5\cdot 0,5 + 2\cdot 8 + 5\cdot 10 + 8,5\cdot 8 + 12,5\cdot 9 + 17,5\cdot 12 + 25\cdot 8}{60} = \frac{659}{60} = 10,983\left[Jahre\right]$$

b)

$$s^2 = \frac{1}{59}\sum_{i=1}^{7}\left(x_i - \bar{x}\right)^2 \cdot n_i$$

$$s^2 = \frac{1}{59}\left[\begin{array}{l}(0,5-10,98)^2\cdot 5 + (2-10,98)^2\cdot 8 + (5-10,98)^2\cdot 10 + (8,5-10,98)^2\cdot 8 \\ +(12,5-10,98)^2\cdot 9 + (17,5-10,98)^2\cdot 12 + (25-10,98)^2\cdot 8\end{array}\right]$$

$$s^2 = \frac{1}{59}\cdot 3.704,484 = 62,7878 \qquad \Rightarrow \qquad s = \sqrt{62,7878} = 7,924$$

Alternative (nicht erwartungstreu):

$$s^2 = \frac{1}{60}\cdot 3.704,484 = 61,741 \qquad \Rightarrow \qquad s = \sqrt{61,741} = 7,86$$

2.) **Lorenz und Gini: Das Fischparadies**

Im Fischdorado „Angelfrisch" gibt es fünf Fischteiche, die mit unterschiedlich vielen Fischen bestückt sind. Die Verteilung der Fische auf die einzelnen Teiche konnte von der Organisation „fgyb" folgendermaßen geschätzt werden:

Teich	1	2	3	4	5
Anzahl der Fische	200	1.000	300	1.400	100

a) Wie könnte die Organisation „fgyb" Ihrer Ansicht nach heißen?
b) Bestimmen Sie das arithmetische Mittel und den Zentralwert.
c) Zeichnen Sie die Lorenzkurve.
d) Für welche Verteilung wäre der Gini-Koeffizient maximal bzw. minimal?

<u>Lösung:</u>

a) freie Antwort – z.B. **F**ish **G**ives **Y**ou **B**rain

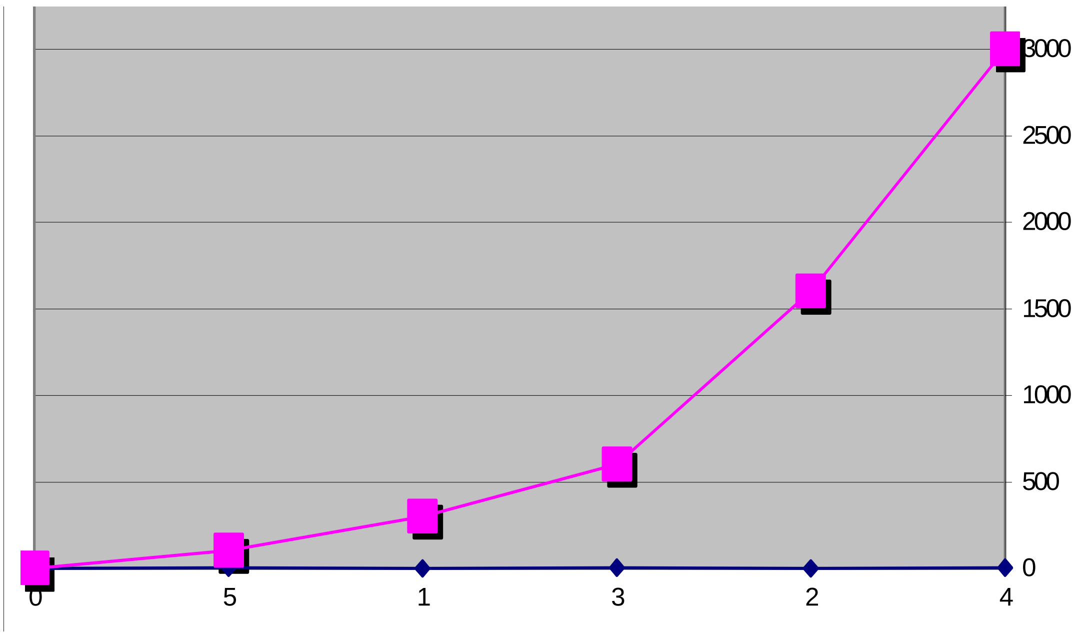
b) neue Anordnung der Größen:

Teich	5	1	3	2	4
Anzahl der Fische	100	200	300	1.000	1.400

Mittelwert: $\bar{x} = \frac{100+200+300+1.000+1.400}{5} = \frac{3.000}{5} = 600\,[Fische]$

Zentralwert ist der mittlere Wert: $\bar{x} = 300\,[Fische]$

c)



d) Der Gini-Koeffizient wäre maximal, wenn sich alle Fische in einem Teich aufhalten würden, hingegen wäre er minimal (GK = 0), wenn in jeden Teich jeweils 600 Fische wäre, also eine Gleichverteilung vorläge.

3.) **Preisindizes**

Das Unternehmen „Knut & Wicht" stellt ein Produkt in drei verschiedenen Variationen A, B und C her. Die folgende Tabelle zeigt die Entwicklung der Absatzmengen und Stückpreise:

Jahr / Var.	2003		2004		2005	
	Preis	**Menge**	**Preis**	**Menge**	**Preis**	**Menge**
A	10	150	12	100	12	100
B	8	200	10	150	10	220
C	12	250	12	250	11	230

a) Bestimmen Sie die Preisindizes nach Laspeyres zur Basis 2003.

b) Bestimmen Sie die Preisindizes nach Paasche zur Basis 2003.

c) Begründen Sie welcher Preisindex aufgrund des vorgegebenen Datenmaterials die Situation treffender charakterisiert.

<u>Lösung:</u>

a) Preisindizes nach Laspeyres mit Basisjahr 2003:

allgemein: $$P_{L\,(0\,;\,t)} = \frac{\sum_{i=1}^{n} p_t^i \cdot q_0^i}{\sum_{i=1}^{n} p_0^i \cdot q_0^i}$$

$$P_{L\,(2003;2004)} = \frac{12 \cdot 150 + 10 \cdot 200 + 12 \cdot 250}{10 \cdot 150 + 8 \cdot 200 + 12 \cdot 250} = \frac{6.800}{6.100} = 1{,}115$$

$$P_{L\,(2003;2005)} = \frac{12 \cdot 150 + 10 \cdot 200 + 11 \cdot 250}{10 \cdot 150 + 8 \cdot 200 + 12 \cdot 250} = \frac{6.550}{6.100} = 1{,}074$$

b) Preisindizes nach Paasche mit Basisjahr 2003:

allgemein:
$$P_{P\,(0\,;\,t)} = \frac{\sum_{i=1}^{n} p_t^i \cdot q_t^i}{\sum_{i=1}^{n} p_0^i \cdot q_t^i}$$

$$P_{P\,(2003;2004)} = \frac{12\cdot 100 + 10\cdot 150 + 12\cdot 250}{10\cdot 100 + 8\cdot 150 + 12\cdot 250} = \frac{5.700}{5.200} = 1,096$$

$$P_{P\,(2003;2005)} = \frac{12\cdot 100 + 10\cdot 220 + 11\cdot 230}{10\cdot 100 + 8\cdot 220 + 12\cdot 230} = \frac{5.930}{5.520} = 1,074$$

c) Der Preisindexanstieg nach Paasche ist nicht so extrem wie der nach Laspeyeres, weil bei Paasche die mengenmäßige Reduzierung der teurer werdenden Produkte mit in die Berechnung eingeht.

Daher wäre der Paasche-Preisindex für die gegebene Situation zutreffender.

4.) **Regressionsgerade und Korrelation**
Ein großer Kaufhauskonzern testet bei fünf seiner Filialen den Einfluss des Preises X (in €) eine Ware auf dessen Absatzmenge Y (in 1.000 Stück). Dafür werden die Preis unterschiedlich hoch festgesetzt. Folgendes Käuferverhalten ist zu beobachten:

Filiale	Preis X	Absatz Y
1	5	16
2	6	15
3	7	12
4	8	9
5	14	8

a) Zeichnen Sie die Punkte in ein Koordinatensystem, bestimmen Sie die Gleichung der linearen Regressionsgeraden und zeichnen Sie diese ebenfalls in das Koordinatensystem.

b) Bestimmen Sie den Korrelationskoeffizienten r zwischen X und Y.

c) Welche Absatzmenge würde der Konzern wohl bei einem Preis von 4 € erwarten?

d) Welches ökonomische Gesetz kann man hier beobachten?

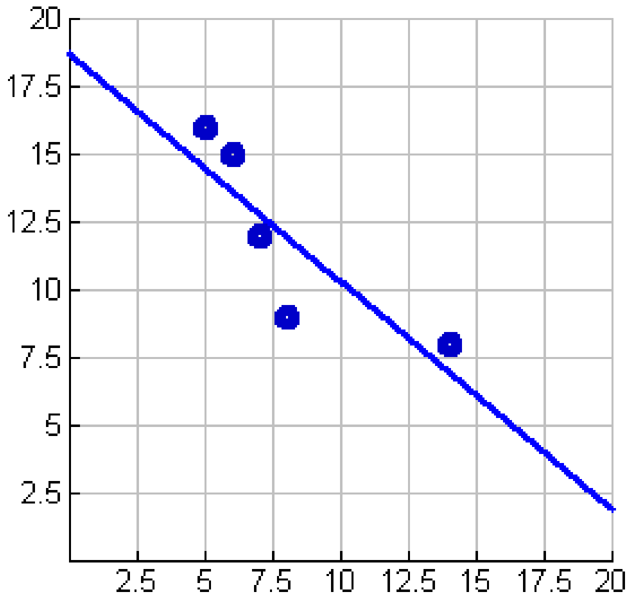
Lösung:

a) Regressionsgerade

Filiale	Preis X	Absatz-menge Y	$x_i - x_{MW}$	$(x_i - x_{MW})^2$	$y_i - y_{MW}$	$(y_i - y_{MW})^2$	$x_i - x_{MW}$ * $y_i - y_{MW}$
1	5	16	-3	9	4	16	-12
2	6	15	-2	4	3	9	-6
3	7	12	-1	1	0	0	0
4	8	9	0	0	-3	9	0
5	14	8	6	36	-4	16	-24
Summe:	40	60	0	50	0	50	-42
Mittelwert:	8	12					

$$b_2 = \frac{-42}{50} = -0,84 \quad \Rightarrow \quad b_1 = 12 - (-0,84) \cdot 8 \quad \Rightarrow \quad b_1 = 18,72$$

$$\Rightarrow \quad y = -0,84x + 18,72$$



b) Korrelationskoeffizient:

$$r = \frac{\sum_{i=1}^{5} \left[(x_i - \mu_x)(y_i - \mu_y) \right]}{\sqrt{\sum_{i=1}^{5} (x_i - \mu_x)^2 \cdot \sum_{i=1}^{5} (y_i - \mu_y)^2}} \quad \Rightarrow \quad r = \frac{-42}{\sqrt{50 \cdot 50}} = -0,84$$

c)

$$f(x) = -0,84x + 18,72$$

$$f(4) = (-0,84) \cdot 4 + 18,72 = 15,36$$

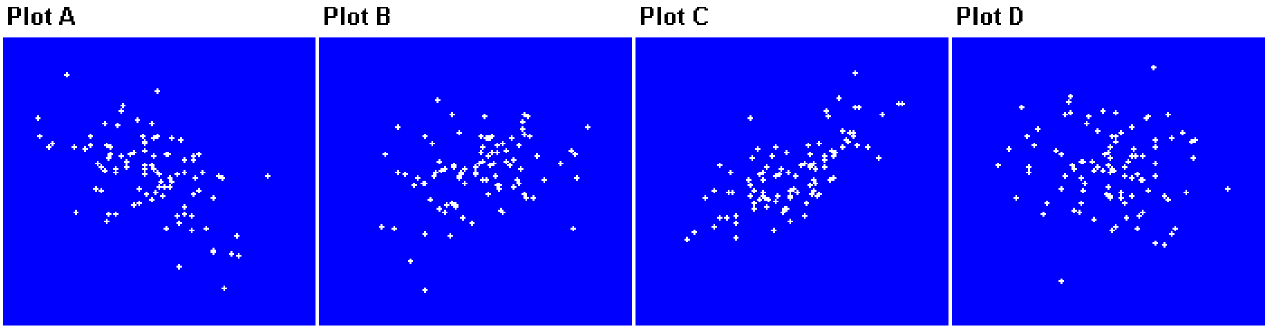
d) Gesetz der Nachfrage:

Mit steigenden Preisen nimmt die Nachfragemenge ab,
mit fallenden Preisen nimmt die Nachfragemenge zu.

5.) **Korrelationschaos**

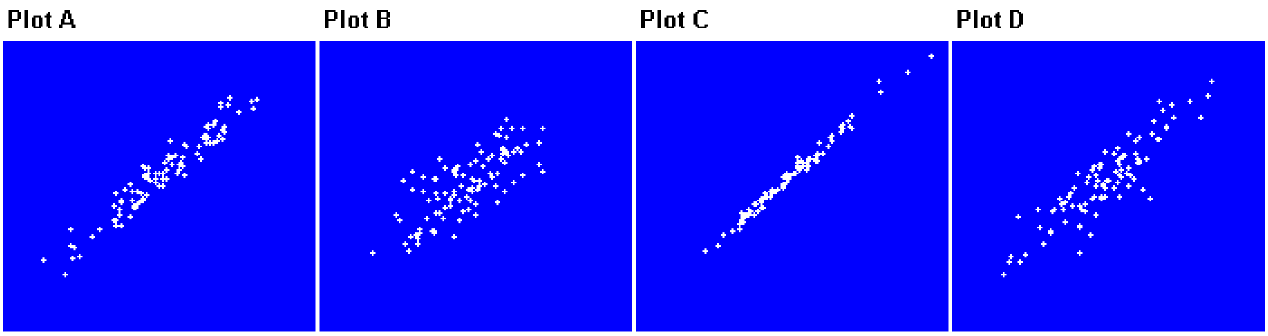
a) Ordnen Sie die Korrelationswerte den Zeichnungen durch Ankreuzen zu:

Teil I:



r = -0.51	○ A	○ B	○ C	○ D
r = -0.09	○ A	○ B	○ C	○ D
r = 0.30	○ A	○ B	○ C	○ D
r = 0.70	○ A	○ B	○ C	○ D

Teil II:



r = 0.74	○ A	○ B	○ C	○ D
r = 0.89	○ A	○ B	○ C	○ D
r = 0.95	○ A	○ B	○ C	○ D
r = 0.98	○ A	○ B	○ C	○ D

b) In welchen Grenzen können sich die Korrelationswerte bewegen?

c) Erklären Sie das Korrelationsverhalten, wenn r die jeweiligen Außengrenzen bzw. Grenzwerte annimmt.

<u>Lösung:</u>

a)

Teil I:

r = -0.51	● A	○ B	○ C	○ D
r = -0.09	○ A	○ B	○ C	● D
r = 0.30	○ A	● B	○ C	○ D
r = 0.70	○ A	○ B	● C	○ D

Teil II:

r = 0.74	○ A	● B	○ C	○ D
r = 0.89	○ A	○ B	○ C	● D
r = 0.95	● A	○ B	○ C	○ D
r = 0.98	○ A	○ B	● C	○ D

b) Die Korrelationswerte können sich in den Grenzen von –1 bis 1 bewegen.

c) Bei r = -1 spricht von negativer Korrelation bzw. entgegengesetztem Verhalten der beiden Merkmalsausprägungen, während r = 1 eine positive Korrelation signalisiert mit gleichlaufendem Verhalten der beiden Merkmalsausprägungen.

6.) **Rangkorrelation: Erst- und Zweitkorrektur**

Die Klausur „Allgemein BWL" wird von zwei Dozenten unabhängig voneinander korrigiert. Um die Übereinstimmung der Bewertung der Korrektoren zu prüfen möchte man den Rangkorrelationskoeffizient für acht Studenten bestimmen.
Folgende Rohpunkte wurde erreicht:

Schüler	A	B	C	D	E	F	G	H
Rohpunkte Korr. 1	98	95	90	84	78	76	68	63
Rohpunkte Korr. 2	94	92	95	85	70	80	75	65

Beurteilen Sie die Qualität der Korrelation der Erst- und Zweitkorrektur auf der Basis geeigneter Rechenoperationen.

Lösung:

Ermittlung der Rangfolge der beiden Korrekturen:

Schüler	A	B	C	D	E	F	G	H
Rohpunkte Korr. 1	1	2	3	4	5	6	7	8
Rohpunkte Korr. 2	2	3	1	4	7	5	6	8

Schüler	x_i	y_i	d_i	d_i^2
A	1	2	-1	1
B	2	3	-1	1
C	3	1	2	4
D	4	4	0	0
E	5	7	-2	4
F	6	5	1	1
G	7	6	1	1
H	8	8	0	0
Summe:				12

Rangkorrelation: $R = 1 - \frac{6 \cdot 12}{8 \cdot 63} = 0,86$

Stochastik

1.) **Binomialverteilte Zufallsvariablen**

Eine Zufallsvariable ist $B_{100;\frac{1}{3}}$ - verteilt.

Bestimmen Sie alle Werte von X, die für **k = 1 und 2** im Intervall $\left| X - \mu \right| \leq k \cdot \sigma$ liegen.

Lösung:

$$B_{100;\frac{1}{3}}: \Rightarrow \mu = 33\frac{1}{3} \Rightarrow \sigma^2 = 22\frac{2}{9} \Rightarrow \sigma = \sqrt{22\frac{2}{9}} = 4{,}714$$

$$Intervall\,1: \left| X - 33\frac{1}{3} \right| \leq 4{,}7 \Rightarrow X \in \{29; 30; \ldots; 37; 38\}$$

$$Intervall\,2: \left| X - 33\frac{1}{3} \right| \leq 9{,}4 \Rightarrow X \in \{24; 25; \ldots; 41; 42\}$$

2.) **Erwartungswert**

In einem ZE gebe die Zufallsvariable X die Anzahl der Treffer an.
Für die Wahrscheinlichkeitsverteilung von X gilt:

k	0	1	2	3	4	sonst
P (X = k)	$\frac{1}{20}a$	$a^2 - 3{,}7$	$\frac{1}{5}a$	$\frac{1}{20}a$	$\frac{1}{10}$	0

Bestimmen Sie a so, dass eine Wahrscheinlichkeitsverteilung vorliegt.

Lösung:

$$\frac{1}{20}a + a^2 - 3{,}7 + \frac{1}{5}a + \frac{1}{20}a + \frac{1}{10} = 1 \Rightarrow a^2 + \frac{3}{10}a - \frac{46}{10} = 0$$

$$\Rightarrow a_1 = 2 \wedge a_2 = -2{,}6$$

Nur a_1 ist als Lösung relevant.

3.) **Die hypergeometrische Lostrommel**

In einer Lostrommel befinden sich 8 Gewinnlose und 42 Nieten.
Rudi Glückspilz kauft 3 Lose.

a) Wie groß ist die Wahrscheinlichkeit,
(i) genau ein Gewinnlos zu ziehen?
(ii) höchstens ein Gewinnlos zu ziehen?

b) Die Gewinnlose unterscheiden sich in zwei Kategorien:
3 Lose von Typ A und 5 Lose von Typ B.
Rudi hat genau ein Gewinnlos gezogen.
Wie hoch ist die Wahrscheinlichkeit, dass Rudi ein Los vom Typ A gezogen hat?

<u>Lösung:</u>

$$a)\quad (i)\quad H(X=1) = \frac{\binom{8}{1}\binom{42}{2}}{\binom{50}{3}} = 0{,}3514 \approx 35{,}14\,\%$$

$$(ii)\quad H(X\leq 1) = H(X=0)+H(X=1)$$

$$H(X\leq 1) = \frac{\binom{8}{0}\binom{42}{3}}{\binom{50}{3}}+0{,}3514 = 0{,}5857+0{,}3514 = 0{,}9371$$

$$b)\quad P_A(B) = \frac{\dfrac{\binom{3}{1}\binom{5}{0}\binom{42}{2}}{\binom{50}{3}}}{\dfrac{\binom{8}{1}\binom{42}{2}}{\binom{50}{3}}} = \frac{\binom{3}{1}}{\binom{8}{1}} = 0{,}375 = 37{,}5\,\%$$

4.) **Die Binomial-Weisen**

In einem fernen Land beurteilen „fünf Weise“ unabhängig voneinander jedes Jahr die zukünftige wirtschaftliche Entwicklung.

Wie groß müsste die Wahrscheinlichkeit p eines richtigen Urteils bei jedem der fünf Experten mind. sein, damit man mit einer Wahrscheinlichkeit von mehr als 90 % nur richtige Vorhersagen erhält?

Lösung:

$$B(X=5) \quad = \quad \binom{5}{5} p^5 \cdot (1-p)^0 \quad \geq \quad 0{,}9$$

$$\Rightarrow \quad p^5 \quad \geq \quad 0{,}9 \quad \xrightarrow{\sqrt[5]{p}} \quad p \quad = \quad 0{,}9791 \quad = \quad 97{,}91\,\%$$

5.) **Poisson-/Binomialverteilung: Fahrradschläuche**

Bei der Produktion von Schläuchen für Fahrräder sind erfahrungsgemäß 5 % defekt.
Für Großhändler werden Packungen zu je 100 Schläuchen geliefert.

Mit welcher Wahrscheinlichkeit befinden sich in einer Packung

a) fünf defekte Schläuche? (=> Poissonverteilung anwenden!)

b) mind. ein defekter Schlauch? (=> Poissonverteilung anwenden!)

c) unter 100 Schläuchen mind. 5 und höchstens 10 defekte? (=> Binomialverteilung anwenden!)

Lösung:

$$P(X=k) \quad = \quad \frac{\mu^k}{k!} e^{-\mu}$$

$$a) \quad P(X=5) \quad = \quad \frac{5^5}{5!} e^{-5} \quad = \quad 0{,}17546 \quad \approx \quad 17{,}55\,\%$$

$$b) \quad P(X \geq 1) \quad = \quad 1 - P(X=0) \quad = \quad 1 - \frac{5^0}{0!} e^{-5} \quad = \quad 0{,}9933 \quad \approx \quad 99{,}33\,\%$$

$$c) \quad B(5 \leq X \leq 10) \quad = \quad B(X \leq 10) - B(X \leq 4) \qquad [n = 100]$$

$$B(5 \leq X \leq 10) \quad = \quad 0{,}9885 - 0{,}4360$$

$$B(5 \leq X \leq 10) \quad = \quad 0{,}5525 \quad = \quad 55{,}25\,\%$$

6.) **Normalverteilung bei Niederschlägen**

Die Zufallsvariable X gebe die Niederschlagsmenge von Ludwigshafen für den Monat März an. Langjährige Aufzeichnungen ergaben, dass die Niederschlagsmenge normalverteilt ist mit dem Mittelwert μ = 75 mm und σ = 5 mm.

Wie groß ist die Wahrscheinlichkeit, dass im kommenden März

a) mindestens 80 mm Niederschlag fallen?

b) höchstens 65 mm Niederschlag fallen?

c) zwischen 53 mm und 85 mm Niederschlag fallen?

d) Wie hoch sind die Niederschlagsmengen für den kommenden März, wenn die Wahrscheinlichkeit bei mind. 99 % im Intervall um den erwarteten Wert liegt?

Lösung:

a)

$$P(X \geq 80) = 1 - P(X < 80) = 1 - \Phi(1) = 1 - 0{,}8413 = 0{,}1587$$

$$Umrechnung: \quad x = \frac{80-75}{5} = 1$$

b)

$$P(X \leq 65) = \Phi(-2) = 1 - \Phi(2) = 1 - 0{,}9772 = 0{,}0228$$

$$Umrechnung: \quad x = \frac{65-75}{5} = -2$$

c)

$$P(53 \leq X \leq 85) = \Phi(2) - \Phi(-4{,}4) = \Phi(2) - \left[1 - \Phi(4{,}4)\right]$$

$$P(53 \leq X \leq 85) = \Phi(2) - 1 + \Phi(4{,}4)$$

$$P(53 \leq X \leq 85) = 0{,}9772 - 1 + 0{,}9999 = 0{,}9771$$

$$Umrechnung: \quad x_1 = \frac{53-75}{5} = -4{,}4 \quad \wedge \quad x_2 = \frac{85-75}{5} = 2$$

d)

$$P(k_1 \leq X \leq k_2) \geq 0{,}99$$

$$\Rightarrow \quad P(k_1 \leq X \leq k_2) = \Phi(2{,}58) - \Phi(-2{,}58) = \Phi(2{,}58) - \left[1 - \Phi(2{,}58)\right]$$

$$\Rightarrow \quad P(k_1 \leq X \leq k_2) = 2 \cdot \Phi(2{,}58) - 1$$

$$Umrechnung\ k_1: \quad -2{,}58 = \frac{k_1 - 75}{5} \quad \Rightarrow \quad k_1 = 62{,}1 \approx 62$$

$$Umrechnung\ k_2: \quad 2{,}58 = \frac{k_2 - 75}{5} \quad \Rightarrow \quad k_2 = 87{,}9 \approx 88$$

7.) **Gurtmuffel & Co.**

Polizeilichen Statistiken zufolge beträgt der Anteil der Autolenker, die während der Fahrt keinen Sicherheitsgurt tragen, 20 %. Diese Fahrer werden ab jetzt „Gurtmuffel" genannt. Man darf annehmen, dass die Autofahrer unabhängig voneinander den Gurt anlegen oder nicht.

a) Wie viele Autos muss man überprüfen, um mit mind. 99,9 %iger Wahrscheinlichkeit mind. einen Gurtmuffel zu finden?

b) Wie groß wäre der Anteil p der Gurtmuffel mind., wenn von 30 vorbeifahrenden Autos mit 95 %iger Wahrscheinlichkeit mind. eines von einem Gurtmuffel gelenkt würde.

<u>Lösung:</u>

$$a)\quad B(X \geq 1) \;=\; 1 - B(X = 0) \;\geq\; 0{,}999$$

$$\Rightarrow \quad \binom{n}{0} \cdot 0{,}2^0 \cdot 0{,}8^n \;\leq\; 0{,}001$$

$$\xrightarrow{\ln} \quad n \cdot \ln(0{,}8) \;\leq\; \ln(0{,}001)$$

$$\Rightarrow \quad n \;\geq\; \frac{\ln(0{,}001)}{\ln(0{,}8)} \;=\; 30{,}96$$

$$\Rightarrow \quad n \;\geq\; 31$$

$$b)\quad B(X \geq 1) \;=\; 1 - B(X = 0) \;\geq\; 0{,}95$$

$$\xrightarrow{\text{vereinfachen}} \quad \binom{30}{0} p^0 \cdot (1-p)^{30} \;\leq\; 0{,}05$$

$$\xrightarrow{\text{vereinfachen und } \sqrt[30]{(1\text{-p})}} \quad (1-p) \;\leq\; \sqrt[30]{0{,}05}$$

$$\xrightarrow{\text{vereinfachen}} \quad p \;\geq\; 1 - \sqrt[30]{0{,}05} \;=\; 0{,}095 \;=\; 9{,}5\,\%$$