

2. Vordiplom: Statistik (2 Stunden)

2. (7 Punkte) Neu entwickelte Mountain-Bike-Felgen mit Keramikbelag sollen durch Staub und Schmutz beim Bremsen weniger rasch abgenutzt werden als die herkömmlichen Aluminium-Felgen. Velohändler Ferdj möchte dies testen und notiert sich die Lebensdauer (in km) der Felgen verschiedener Kunden.

Daten:

Alu	3662	4247	3920	3673	4047	4091	4676	3966
Keramik	4386	4089	4736	4348	4414	4995		

Kennzahlen:

Alu	$\bar{x} = 4035.25$	$s_x^2 = 106'752$
Keramik	$\bar{y} = 4494.67$	$s_y^2 = 102'594$

- 1) Hier wird an jeder Versuchseinheit
 - a) mehr als eine Messung gemacht.
 - b) genau eine Messung gemacht.
- 2) Wie lautet die korrekte Nullhypothese?
 - a) Keramikelgen haben eine längere Lebensdauer als Alufelgen.
 - b) Es gibt einen Unterschied zwischen der Lebensdauer von Keramikelgen und Alufelgen.
 - c) Die Lebensdauer ist für Keramikelgen und Alufelgen identisch.
- 3) Wie lautet die korrekte Alternativhypothese?
 - a) Keramikelgen haben eine längere Lebensdauer als Alufelgen.
 - b) Alufelgen haben eine längere Lebensdauer als Keramikelgen.
 - c) Es gibt einen Unterschied zwischen der Lebensdauer von Keramikelgen und Alufelgen.
 - d) Die Lebensdauer ist für Keramikelgen und Alufelgen identisch.

Viel Glück!

1. (8 Punkte)

Das Wundermittel NIE-PLAT kann in einen platten Reifen gespritzt werden und dichtet diesen ab. Gemäss dem Hersteller führt das Mittel in 75% der Fälle zum Erfolg.

a) 100 platte Reifen werden mit NIE-PLAT behandelt. Wie gross ist der Erwartungswert der Anzahl erfolgreich reparierter Reifen?

An einem grossen Volks-Radrennen wird das Mittel gratis verteilt, danach werden alle Personen befragt, die es benutzt haben. Es zeigte sich, dass 73 von 100 Anwendern erfolgreich waren.

b) Schätzen Sie aufgrund dieser Beobachtung die Wahrscheinlichkeit, dass man mit NIE-PLAT erfolgreich eine Reifenpanne beheben kann.

c) Berechnen Sie (mit einer geeigneten Approximation) die Wahrscheinlichkeit, dass man, unter der Annahme einer Erfolgswahrscheinlichkeit von 75%, nur 73 oder weniger erfolgreiche Reparaturen beobachtet hat.

Hinweis:

$$\frac{\sqrt{0.75 \cdot 0.25 \cdot 100}}{1} = \frac{\sqrt{\frac{16}{300}}}{1} \approx 0.23.$$

d) Sie sind misstrauisch und denken, dass die wahre Erfolgswahrscheinlichkeit tiefer liegt als die Herstellerangabe von 75%. Führen Sie mit Hilfe des Resultates im Teilaufgabe c) einen statistischen Test auf dem 5% Niveau durch, ob die Aussage des Herstellers haltbar ist. Geben Sie die Nullhypothese, die Alternative und die Testentscheidung an.

e) Geben Sie ein approximatives 95%-Vertrauensintervall für die Erfolgswahrscheinlichkeit von NIE-PLAT an.

Hinweis:

$$\sqrt{0.73 \cdot 0.27 \cdot 100} \approx 4.43.$$

- 5) Wenn die beobachtete Varianz der Lebensdauer für Aluminium- und Keramikelgen kleiner wären (bei gleichbleibenden Mittelwerten), dann wäre der *p*-Wert...
 - a) grösser.
 - b) kleiner.
 - c) mit den vorliegenden Angaben ist keine solche Aussage über den *p*-Wert möglich.

4) Wenn wir für Aluminium- und Keramikelgen je doppelt so viele Beobachtungen hätten, Mittelwerte und Streuungen sich aber nicht ändern würden, dann wäre der *p*-Wert...

- a) grösser.
- b) kleiner.
- c) gleichbleibend.

5) Wenn die beobachtete Varianz der Lebensdauer für Aluminium- und Keramikelgen kleiner wären (bei gleichbleibenden Mittelwerten), dann wäre der *p*-Wert...

Nimm jetzt an, der korrekt durchgeführte Test ergibt einen *p*-Wert von 0.011, bzw. 1.1%.

- d) Die Lebensdauer ist für Keramikelgen und Alufelgen identisch.
- c) Es gibt einen Unterschied zwischen der Lebensdauer von Keramikelgen und Alufelgen.
- b) Alufelgen haben eine längere Lebensdauer als Keramikelgen.
- a) Keramikelgen haben eine längere Lebensdauer als Alufelgen.

Wie lautet die korrekte Nullhypothese?

- a) mehr als eine Messung gemacht.
- b) genau eine Messung gemacht.

Hier wird an jeder Versuchseinheit

- a) Keramikelgen haben eine längere Lebensdauer als Alufelgen.
- b) Es gibt einen Unterschied zwischen der Lebensdauer von Keramikelgen und Alufelgen.
- c) Die Lebensdauer ist für Keramikelgen und Alufelgen identisch.

Wie lautet die korrekte Alternativhypothese?

- a) Keramikelgen haben eine längere Lebensdauer als Alufelgen.
- b) Es gibt einen Unterschied zwischen der Lebensdauer von Keramikelgen und Alufelgen.
- c) Die Lebensdauer ist für Keramikelgen und Alufelgen identisch.

Wie lautet die korrekte Nullhypothese?

- a) Keramikelgen haben eine längere Lebensdauer als Alufelgen.
- b) Es gibt einen Unterschied zwischen der Lebensdauer von Keramikelgen und Alufelgen.
- c) Die Lebensdauer ist für Keramikelgen und Alufelgen identisch.

Wie lautet die korrekte Alternativhypothese?

- a) Keramikelgen haben eine längere Lebensdauer als Alufelgen.
- b) Alufelgen haben eine längere Lebensdauer als Keramikelgen.
- c) Es gibt einen Unterschied zwischen der Lebensdauer von Keramikelgen und Alufelgen.
- d) Die Lebensdauer ist für Keramikelgen und Alufelgen identisch.

6) Nehme an, ein solcher Test werde auf dem 5%-Niveau durchgeführt: Wie wahrscheinlich ist es, dass man fälschlicherweise die Entscheidung trifft, dass die Laufleistung der Keramikefugen höher ist, obwohl sie in Wahrheit mit derjenigen der Alufelgen identisch ist?

- 7) Welchen Schluss kann man aus dem Test ziehen?
- a) 5%
 - b) 98.9%
 - c) 2.5%
 - d) 1.1%

a) Die Nullhypothese ist statistisch mit einem p-Wert von 1.1% gesichert.

b) Es ist mit einem p-Wert von 1.1% gesichert, dass die Laufleistung von Keramikefugen höher ist als jene von Alufelgen.
 c) Die Alternativhypothese kann mit einer Wahrscheinlichkeit von 1.1% widerlegt werden.
 d) Mit einer Irrtumswahrscheinlichkeit von 98.9% haben wir bewiesen, dass Alufelgen besser als Keramikefugen sind.

3. (7 Punkte)

Die Gesundheitskosten (in Franken pro Jahr) pro Person werden in einem einfachen Modell durch die Anzahl Ärzte pro 1000 Einwohner an ihrem Wohnort und durch ihr Alter erklärt.

Das Regressionsmodell lautet:

$$\text{Gesundheitskosten}_i = \alpha + \beta_1 \cdot \text{Ärztedichte}_i + \beta_2 \cdot \text{Alter}_i + E_i, \quad E_i \sim \mathcal{N}(0, \sigma^2)$$

Der Regressionsoutput sieht folgendermassen aus:

Coefficients:

(Intercept)	Estimate	Std. Error	t value	Pr(> t)
aerzte	310.26	128.38	2.417	0.03882 *
alter	19.00	4.13	4.600	...

Residual standard error: 149.9 on 9 degrees of freedom

1) Mit wievielen Beobachtungen wurde diese Regression gerechnet?

- a) 8
- b) 9
- c) 10
- d) 11
- e) 12

2) Wie gross ist die Residuenquadratsumme dieser Regression?

- a) 1'349.1
- b) 149.9
- c) 202'230.09
- d) 325.7

3) Ist es statistisch gesichert, dass ein Zusammenhang zwischen dem Alter einer Person und der Höhe ihrer Gesundheitskosten besteht?

- a) ja
- b) nein
- c) kann mit den vorliegenden Angaben nicht entschieden werden

4) Wie hoch sind die prognostizierten Gesundheitskosten für eine Person im Alter von 50 Jahren, die an einem Ort mit Ärztedichte 1.0 pro 1000 Einwohner zuhause ist?

- a) 950.00
- b) 1260.26
- c) 1333.69
- d) 1296.69

4. (5 Punkte)

X_1 und X_2 sind zwei unabhängige Zufallsvariablen. Beide sind normalverteilt mit gleichem Mittelwert μ , aber unterschiedlicher Standardabweichung σ_1 und σ_2 .

$$X_1 \sim \mathcal{N}(\mu, \sigma_1^2), \quad X_2 \sim \mathcal{N}(\mu, \sigma_2^2).$$

Wir bilden eine Linearkombination

$$X = a \cdot X_1 + b \cdot X_2.$$

a) Berechnen Sie den Erwartungswert $E[X]$ in Abhängigkeit von a und b . Unter welcher Bedingung an a und b ist $E[X] = \mu$?

b) Berechnen Sie die Varianz von X in Abhängigkeit von a und b .
 c) Bestimmen Sie a und b so, dass $E[X] = \mu$ und $Var[X]$ minimal ist.

5. (7 Punkte)

1) Mit welcher der folgenden Verteilungen lässt sich am besten die Anzahl Moleküle CO_2 pro Kubikmeter Luft beschreiben?

- a) Binomial
- b) Poisson
- c) Exponential
- d) uniforme Verteilung

2) Mit welcher der folgenden Verteilungen lässt sich am besten die Länge des Wartens an einer Strassenkreuzung beschreiben?

- a) Binomial
- b) Poisson
- c) Exponential
- d) uniforme Verteilung

5) Welches Intervall ist ein exaktes 95%-Vertrauensintervalle für β_2 ?

- a) $19.00 \pm 2.26 \cdot 4.13$
- b) $19.00 \pm 1.96 \cdot 4.13$
- c) $19.00 \pm 2.26 \cdot \frac{4.13}{\sqrt{6}}$
- d) $19.00 \pm 4.60 \cdot \frac{4.13}{\sqrt{6}}$

6) Kann man aufgrund des p-Werts von 0.03882 für β_1 entscheiden, ob das 95%-Vertrauensintervall für β_1 den Wert 0 enthält?

- a) ja, denn das 95%-Vertrauensintervall enthält immer den Wert 0.
- b) nein, das Resultat aus dem Signifikanztest reicht für diese Aussage nicht.
- c) ja, diese Entscheidung ist generell möglich.
- d) man kann dies nur entscheiden, falls der Test verworfen wurde.

7) Es wird eine einfache lineare Regression mit nur noch der Ärztedichte als erklärender Variable durchgeführt. Ist der Einfluss der Ärztedichte auf die Gesundheitskosten weiterhin signifikant?

- a) auf jeden Fall.
- b) auf keinen Fall.
- c) kann mit den vorliegenden Angaben nicht entschieden werden.

3) Mit welcher der folgenden Verteilungen lässt sich am besten die Anzahl Velos mit grüner Lackierung unter 10 zufällig herausgesuchten Velos beschreiben?

- a) Binomial
- b) Poisson
- c) Exponential
- d) uniforme Verteilung

4) Mit welcher der folgenden Verteilungen lässt sich am besten die Anzahl Telefonanrufe beschreiben, die pro Stunde bei einer Telefonzentrale eingingen?

- a) Binomial
- b) Poisson
- c) Exponential
- d) uniforme Verteilung

5) Die Funktion $f(x)$ ist gegeben durch

$$f(x) = \begin{cases} 0 & x < 0 \\ a - b \cdot x & 0 \leq x < a/b \\ 0 & x \geq a/b \end{cases}$$

Für welche Parameter a und b kann die Funktion $f(x)$ die Dichte einer

Zufallsvariablen sein?

- a) $a=2, b=3$
- b) $a=1, b=2$
- c) $a=1, b=0$
- d) $a=2, b=2$

e) keine der angegebenen Möglichkeiten

6) Ein fairer Würfel wird 1800 mal geworfen. X sei die Anzahl Würfe mit einer Augenzahl 6. Wie gross ist die Varianz von X ?

- a) 6
- b) 30
- c) 250
- d) 300
- e) 600
- f) 1800

7) Angenommen, wir wollen testen (auf dem 5% Niveau), ob ein Würfel fair ist. Als Teststatistik dient die Anzahl Würfe X mit Augenzahl 6. Der Würfel wird 1800 mal geworfen. Wir beobachten einen Wert $X = 301$. Die folgende Aussage ist dann richtig:

- a) Es ist statistisch gesichert, dass der Würfel unfair ist.
- b) Es ist statistisch gesichert, dass der Würfel fair ist.
- c) Keine der beiden obigen Aussagen.

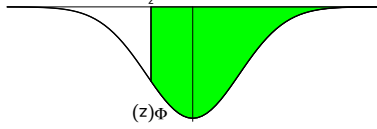
$z \mid$.00 .01 .02 .03 .04 .05 .06 .07 .08 .09

.0	0.5000	0.5040	0.5080	0.5120	0.5160	0.5199	0.5239	0.5279	0.5319	0.5359
.1	0.5398	0.5438	0.5478	0.5517	0.5557	0.5596	0.5636	0.5675	0.5714	0.5753
.2	0.5793	0.5832	0.5871	0.5910	0.5948	0.5987	0.6026	0.6064	0.6103	0.6141
.3	0.6179	0.6217	0.6255	0.6293	0.6331	0.6368	0.6406	0.6443	0.6480	0.6517
.4	0.6554	0.6591	0.6628	0.6664	0.6700	0.6736	0.6772	0.6808	0.6844	0.6879
.5	0.6915	0.6950	0.6985	0.7019	0.7054	0.7088	0.7123	0.7157	0.7190	0.7224
.6	0.7257	0.7291	0.7324	0.7357	0.7389	0.7422	0.7454	0.7486	0.7517	0.7549
.7	0.7580	0.7611	0.7642	0.7673	0.7704	0.7734	0.7764	0.7794	0.7823	0.7852
.8	0.7881	0.7910	0.7939	0.7967	0.7995	0.8023	0.8051	0.8078	0.8106	0.8133
.9	0.8159	0.8186	0.8212	0.8238	0.8264	0.8289	0.8315	0.8340	0.8365	0.8389
1.0	0.8413	0.8438	0.8461	0.8485	0.8508	0.8531	0.8554	0.8577	0.8599	0.8621
1.1	0.8643	0.8665	0.8686	0.8708	0.8729	0.8749	0.8770	0.8790	0.8810	0.8830
1.2	0.8849	0.8869	0.8888	0.8907	0.8925	0.8944	0.8962	0.8980	0.8997	0.9015
1.3	0.9032	0.9049	0.9066	0.9082	0.9099	0.9115	0.9131	0.9147	0.9162	0.9177
1.4	0.9192	0.9207	0.9222	0.9236	0.9251	0.9265	0.9279	0.9292	0.9306	0.9319
1.5	0.9332	0.9345	0.9357	0.9370	0.9382	0.9394	0.9406	0.9418	0.9429	0.9441
1.6	0.9452	0.9463	0.9474	0.9484	0.9495	0.9505	0.9515	0.9525	0.9535	0.9545
1.7	0.9554	0.9564	0.9573	0.9582	0.9591	0.9599	0.9608	0.9616	0.9625	0.9633
1.8	0.9641	0.9649	0.9656	0.9664	0.9671	0.9678	0.9686	0.9693	0.9700	0.9706
1.9	0.9713	0.9719	0.9726	0.9732	0.9738	0.9744	0.9750	0.9756	0.9761	0.9767
2.0	0.9772	0.9778	0.9783	0.9788	0.9793	0.9798	0.9803	0.9808	0.9812	0.9817
2.1	0.9821	0.9826	0.9830	0.9834	0.9838	0.9842	0.9846	0.9850	0.9854	0.9857
2.2	0.9861	0.9864	0.9868	0.9871	0.9875	0.9878	0.9881	0.9884	0.9887	0.9890
2.3	0.9893	0.9896	0.9898	0.9901	0.9904	0.9906	0.9909	0.9911	0.9913	0.9916
2.4	0.9918	0.9920	0.9922	0.9925	0.9927	0.9929	0.9931	0.9932	0.9934	0.9936
2.5	0.9938	0.9940	0.9941	0.9943	0.9945	0.9946	0.9948	0.9949	0.9951	0.9952
2.6	0.9953	0.9955	0.9956	0.9957	0.9959	0.9960	0.9961	0.9962	0.9963	0.9964
2.7	0.9965	0.9966	0.9967	0.9968	0.9969	0.9970	0.9971	0.9972	0.9973	0.9974
2.8	0.9974	0.9975	0.9976	0.9977	0.9977	0.9978	0.9979	0.9979	0.9980	0.9981
2.9	0.9981	0.9982	0.9982	0.9983	0.9984	0.9985	0.9985	0.9986	0.9986	0.9986
3.0	0.9987	0.9987	0.9987	0.9988	0.9988	0.9988	0.9989	0.9989	0.9990	0.9990
3.1	0.9990	0.9991	0.9991	0.9991	0.9991	0.9992	0.9992	0.9992	0.9993	0.9993
3.2	0.9993	0.9994	0.9994	0.9994	0.9994	0.9994	0.9994	0.9995	0.9995	0.9995
3.3	0.9995	0.9996	0.9996	0.9996	0.9996	0.9996	0.9996	0.9996	0.9996	0.9997
3.4	0.9997	0.9997	0.9997	0.9997	0.9997	0.9997	0.9997	0.9997	0.9997	0.9998

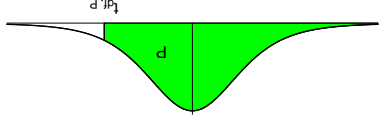
Tabelle der Quadratwurzeln und Quadratzahlen

x^2	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16
$\sqrt{x^2}$	1	1.41	1.73	2	2.24	2.45	2.65	2.83	3	3.16	3.32	3.46	3.61	3.74	3.87	4
x^2	17	18	19	20	25	30	35	40	45	50	55	60	65	70	75	80
$\sqrt{x^2}$	4.12	4.24	4.36	4.47	5	5.48	5.92	6.32	6.71	7.07	7.42	7.75	8.07	8.37	8.64	9
x^2	289	324	361	400	441	484	529	576	625	676	729	784	841	900	961	1000
$\sqrt{x^2}$	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31	32

Tabelle der kumulativen Normalverteilung $\Phi(z) = P[Z \leq z]$, $Z \sim \mathcal{N}(0,1)$



Perzentile der t-Verteilung



Bsp.: $t_{9; 0.975} = 2.262$

df	$t_{0.60}$	$t_{0.70}$	$t_{0.80}$	$t_{0.90}$	$t_{0.95}$	$t_{0.975}$	$t_{0.995}$
1	0.325	0.727	1.376	3.078	6.314	12.706	31.821
2	0.289	0.617	1.061	1.886	2.920	4.303	6.965
3	0.277	0.584	0.978	1.638	2.353	3.182	4.541
4	0.271	0.569	0.941	1.533	2.132	2.776	3.747
5	0.267	0.559	0.920	1.476	2.015	2.571	3.365
6	0.265	0.553	0.906	1.440	1.943	2.447	3.143
7	0.263	0.549	0.896	1.415	1.895	2.365	2.998
8	0.262	0.546	0.889	1.397	1.860	2.306	2.896
9	0.261	0.543	0.883	1.383	1.833	2.262	2.821
10	0.260	0.542	0.879	1.372	1.812	2.228	2.764
11	0.260	0.540	0.876	1.363	1.796	2.201	2.718
12	0.259	0.539	0.873	1.356	1.782	2.179	2.681
13	0.259	0.538	0.870	1.350	1.771	2.160	2.650
14	0.258	0.537	0.868	1.345	1.761	2.145	2.624
15	0.258	0.536	0.866	1.341	1.753	2.131	2.602
16	0.258	0.535	0.865	1.337	1.746	2.120	2.583
17	0.257	0.534	0.863	1.333	1.740	2.110	2.567
18	0.257	0.534	0.862	1.330	1.734	2.101	2.552
19	0.257	0.533	0.861	1.328	1.729	2.093	2.539
20	0.257	0.533	0.860	1.325	1.725	2.086	2.528
21	0.257	0.532	0.859	1.323	1.721	2.080	2.518
22	0.256	0.532	0.858	1.321	1.717	2.074	2.508
23	0.256	0.532	0.858	1.319	1.714	2.069	2.500
24	0.256	0.531	0.857	1.318	1.711	2.064	2.492
25	0.256	0.531	0.856	1.316	1.708	2.060	2.485
26	0.256	0.531	0.856	1.315	1.706	2.056	2.479
27	0.256	0.531	0.855	1.314	1.703	2.052	2.473
28	0.256	0.530	0.855	1.313	1.701	2.048	2.467
29	0.256	0.530	0.854	1.310	1.699	2.045	2.462
30	0.256	0.530	0.854	1.310	1.697	2.042	2.457
40	0.255	0.529	0.851	1.303	1.684	2.021	2.423
60	0.254	0.527	0.848	1.296	1.671	2.000	2.390
120	0.254	0.526	0.845	1.289	1.658	1.980	2.358
∞	0.253	0.524	0.842	1.282	1.645	1.960	2.326