

Statistischer Leitfaden für die Rechtsprechung in Wirtschaftlichkeitsverfahren

Autor

Dr. med. Michel Romanens

Im März 2025

Grundsätzliches zu diesem Manual

Dieser Leitfaden für Juristinnen und Juristen behandelt das Problem des "Patient Profiling" und des "Physician Profiling" anhand von Regressionsanalysen, welche die Krankenversicherer in der Schweiz für drei Controlling-Bereiche verwenden:

- 1) Risikoausgleich unter den Krankenkassen (gemeinsame Einrichtung KVG.org)
- 2) Regressionsindex von Santésuisse,
- 3) Data Mining (Blacklight Analytics).

Zum Leitfaden ist Folgendes festzuhalten

1. Der Leitfaden basiert auf etabliertem statistischem Standardwissen und anerkannten Methoden der Regressionsanalyse, die in der wissenschaftlichen Literatur und Praxis weit verbreitet sind. Beispiel für Statistikprogramm:
<https://www.medcalc.org/manual/regression.php>
2. Die beschriebenen Verfahren entsprechen den tatsächlich von Krankenversicherern eingesetzten Methoden für Risikoausgleich, Regressionsindizes und Data Mining, wie sie in offiziellen Dokumenten und Berichten beschrieben werden.
3. Der Leitfaden stellt keine persönlichen Meinungen dar, sondern fasst objektiv die statistischen Grundlagen zusammen, die für das Verständnis der Profiling-Mechanismen relevant sind.
4. Die Informationen sind kongruent mit dem grundsätzlichen statistischen Vorgehen der Krankenversicherer beim Profiling, auch wenn Details der konkreten Umsetzung abweichen können.
5. Der Leitfaden dient dazu, Juristen ein Grundverständnis der verwendeten Methoden zu vermitteln, nicht um spezifische Verfahren einzelner Versicherer zu bewerten.

Disclaimer

Dieser statistische Leitfaden fasst etablierte Methoden der Regressionsanalyse zusammen, die für das Verständnis von Profiling-Verfahren im Gesundheitswesen relevant sind. Er basiert auf öffentlich zugänglichen Informationen und wissenschaftlicher Literatur. Der Leitfaden erhebt keinen Anspruch auf Vollständigkeit oder detaillierte Beschreibung spezifischer Verfahren einzelner Akteure. Er dient ausschließlich der grundlegenden Information von Juristinnen und Juristen über statistische Konzepte und stellt keine Bewertung oder Kritik konkreter Praktiken dar.

Dokumente zur Profiling-Methode der Krankenversicherer

<https://physicianprofiling.ch/OBSANBerichtArztPruefung102004.pdf>

<https://docfind.ch/Saez2014PCGRomanensWarmuth.pdf>

<https://digitalcollection.zhaw.ch/items/38a04665-340b-4552-a3bb-99ec4bf21969>

<https://docfind.ch//WZWGutachtenWarmuth032016.pdf>

<https://docfind.ch/PhysicianProfilingStatisticPresentation082024.pdf>

<https://docfind.ch/RSAReport2024.pdf>

<https://docfind.ch/VEMSBerichtPCG.pdf>

<https://docfind.ch/VEMSDataMiningReport2024.pdf>

<https://igprof.ch/profiling-risk/>

<https://docfind.ch/references.html>

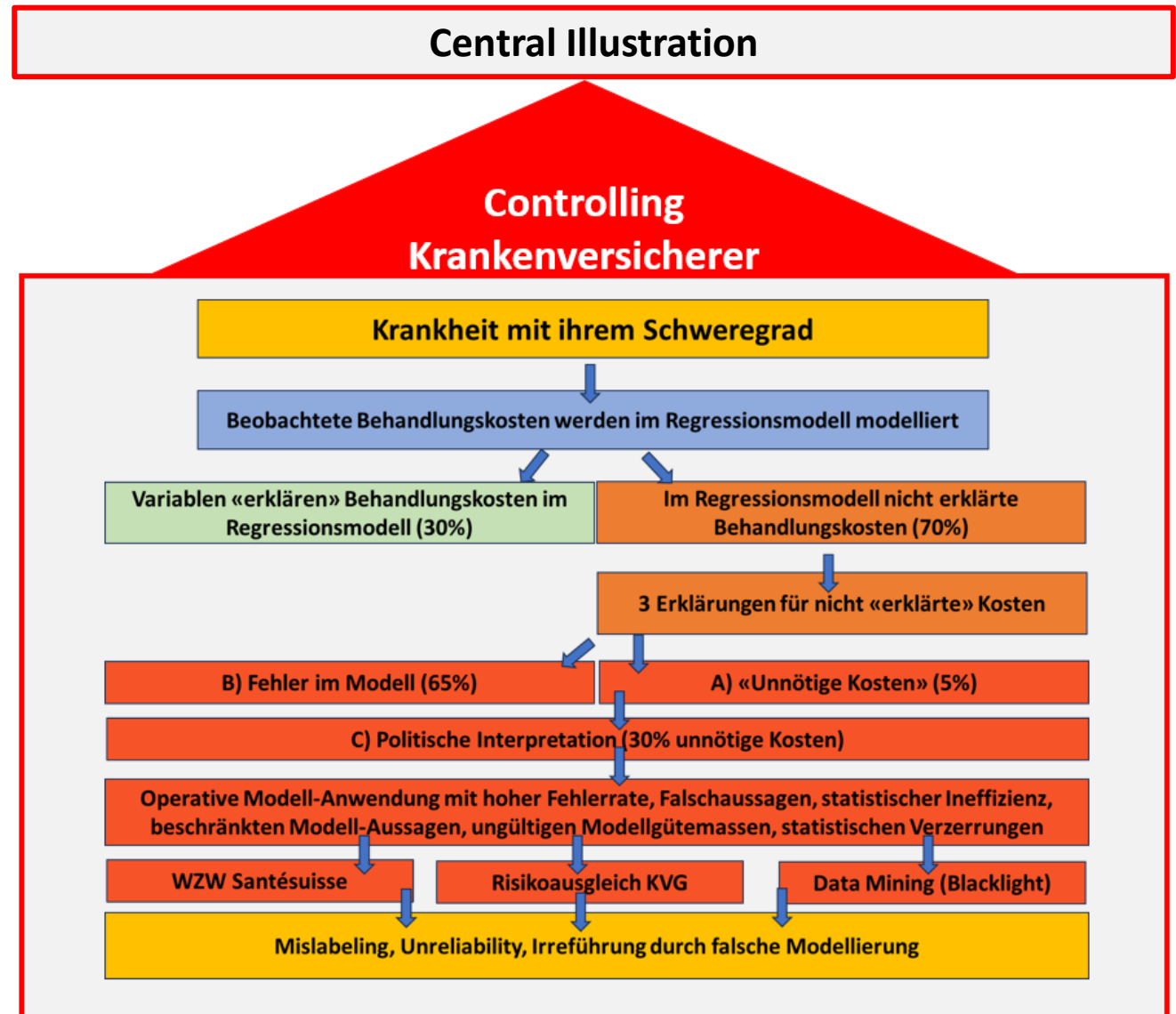
Diese Dokumente im Internet:

Statistischer Leitfaden: Herleitung des Regressionsindex

<https://docfind.ch/StatistischerLeitfaden.pdf>

Statistischer Leitfaden: Beispiele und Berechnungen im Excel Format

<https://docfind.ch/StatistischerLeitfaden.xlsx>



Statistische Grundlagen für Juristinnen und Juristen zu Fragen der Wirtschaftlichkeitsverfahren in der Schweiz

Wirtschaftlichkeitsprüfungen und statistische Methoden

1.1. Beispiel zu Alter und Kosten

In der Regel steigen mit höherem Alter die Kosten (für die medizinische Versorgung).

Betrachten wir nun 20 Personen mit unterschiedlichem Alter.

In diesem fiktiven Beispiel hat jeder Patient eine Nummer (von 1 bis 20).

Die zweite Kolonne zeigt das Alter jedes Patienten.

Patient Nr	Alter
1	30
2	32
3	34
4	36
5	38
6	40
7	42
8	44
9	46
10	48
11	50
12	52
13	54
14	56
15	58
16	60
17	62
18	64
19	66
20	68

Statistische Grundlagen für Juristinnen und Juristen zu Fragen der Wirtschaftlichkeitsverfahren in der Schweiz

Wirtschaftlichkeitsprüfungen und statistische Methoden

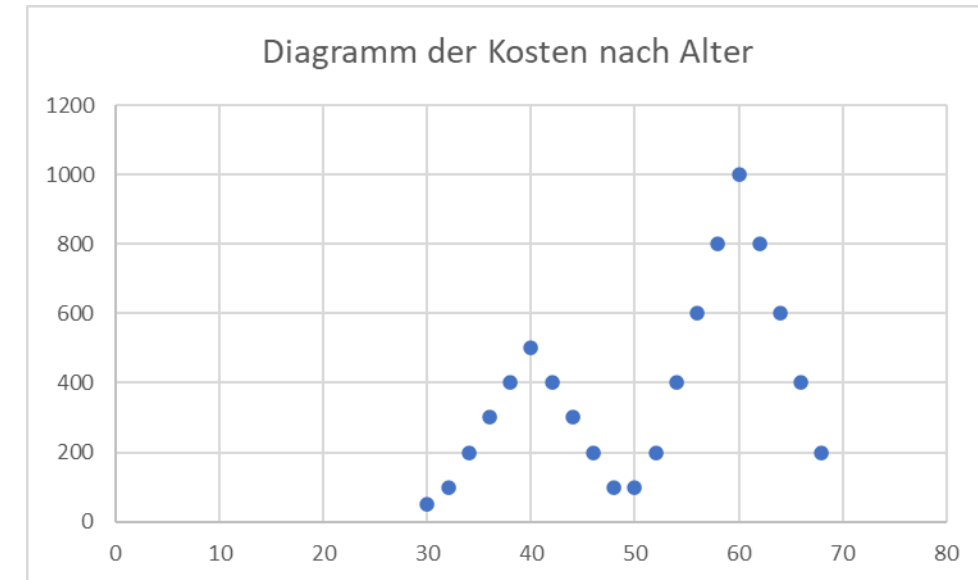
1.2. Beispiel zu Alter und Kosten

Jetzt wird jedem Patienten eine Zahl zu den Kosten zugeordnet.

In der Horizontalen (sogenannte x-Achse) befindet sich das Alter, in der Vertikalen (y-Achse) finden sich die Kosten.

Wir sehen, dass die Kosten zwischen 30-48 ansteigen und dann wieder abfallen, gleiche Beobachtung für 50-68 Jahre).

Patient Nr	Alter	Kosten (CHF)
1	30	50
2	32	100
3	34	200
4	36	300
5	38	400
6	40	500
7	42	400
8	44	300
9	46	200
10	48	100
11	50	100
12	52	200
13	54	400
14	56	600
15	58	800
16	60	1000
17	62	800
18	64	600
19	66	400
20	68	200



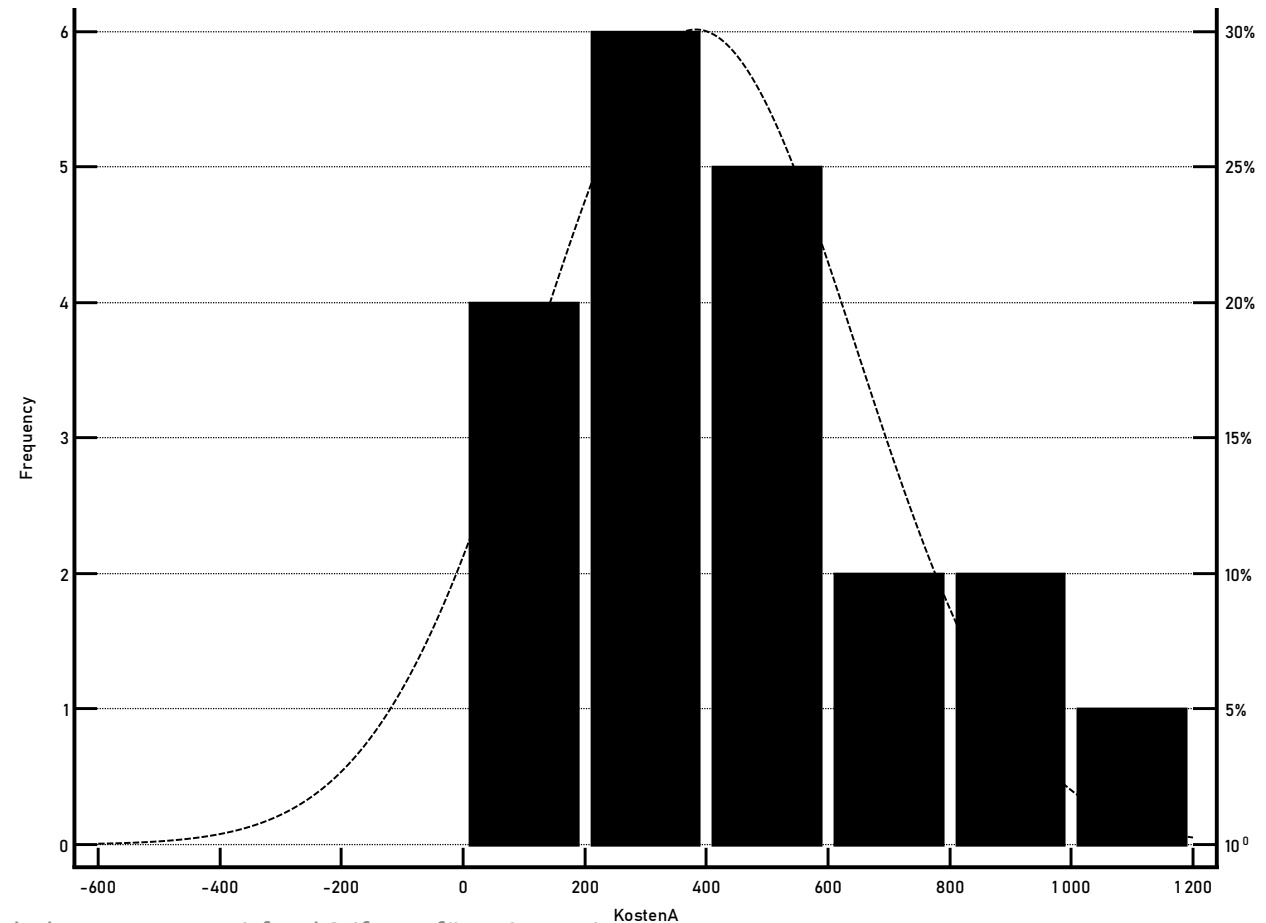
Statistische Grundlagen für Juristinnen und Juristen zu Fragen der Wirtschaftlichkeitsverfahren in der Schweiz

Wirtschaftlichkeitsprüfungen und statistische Methoden

1.3. Beispiel zu Alter und Kosten

Jetzt wollen wir wissen, ob diese Zahlen normal verteilt sind. Die Normal-Verteilung ist eine Voraussetzung für die Regressions-Statistiken. Die gestrichelte Linie ist die Gauss-Kurve, die Balken ist die Summe der einzelnen Kostenwerte.

Das Statistikprogramm MedCalc® zeigt graphisch, dass die Kosten von 20 Personen normal verteilt sind (accept normality $p=0.0792$).



Statistische Grundlagen für Juristinnen und Juristen zu Fragen der Wirtschaftlichkeitsverfahren in der Schweiz

Wirtschaftlichkeitsprüfungen und statistische Methoden

1.4.1. Beispiel zu Alter und Kosten

Santésuisse bildet gerne Klassen, um ihre Statistikberechnungen auf den Erklärungsgehalt zu prüfen.

Jetzt wollen wir wissen, wie die Kosten für 4 Altersklassen verteilt sind.

Altersklassen sollen sein: 30-38, 40-48, 50-58, 60-68.

Wir sehen, dass mit zunehmender höherer Altersklasse der Mittelwert der Kosten ansteigt, von 255 auf 600.

Zahlenreihe A (normal verteilt)

Altersklasse	Alterswert	Mittelwert
30-38	30	255
40-48	40	300
50-58	50	420
60-68	60	600

Statistische Grundlagen für Juristinnen und Juristen zu Fragen der Wirtschaftlichkeitsverfahren in der Schweiz

Wirtschaftlichkeitsprüfungen und statistische Methoden

1.4.2. Beispiel zu Alter und Kosten

In der Graphik sehen wir nun die 4 Altersklassen in der x-Achse und die Durchschnittskosten in der y-Achse.

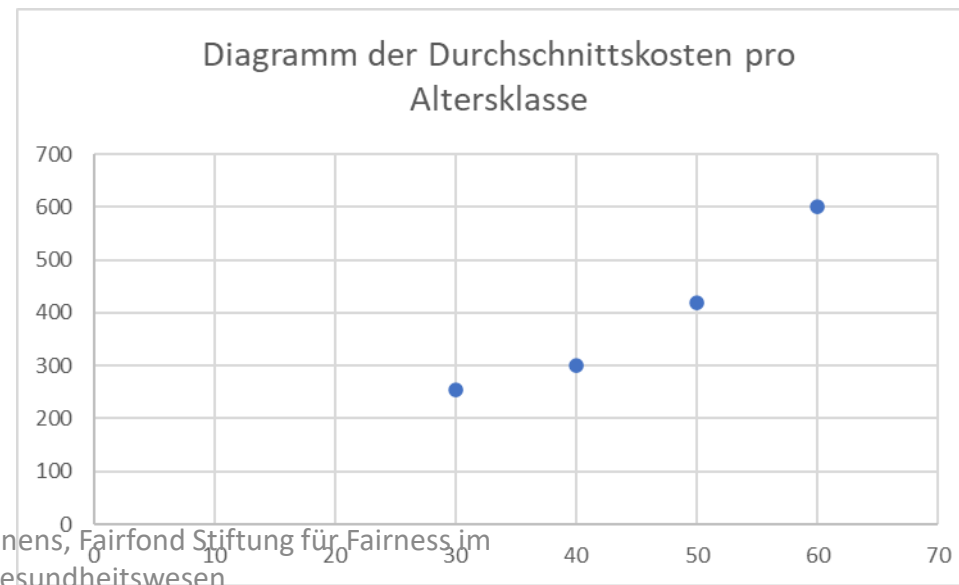
Wir beobachten einen stetigen Anstieg der Kosten pro Altersklasse.

Ergebnis ist also eindeutig! Die Kosten steigen mit höherem Alter? Falsch! (siehe Graphik 1.2).

Es muss heissen: Die Durchschnittskosten steigen mit höherer Altersklasse.

Zahlenreihe A (normal verteilt)

Altersklasse	Alterswert	Mittelwert
30-38	30	255
40-48	40	300
50-58	50	420
60-68	60	600



Statistische Grundlagen für Juristinnen und Juristen zu Fragen der Wirtschaftlichkeitsverfahren in der Schweiz

Wirtschaftlichkeitsprüfungen und statistische Methoden

1.4.3. Beispiel zu Alter und Kosten

Jetzt wollen wir wissen, zu wieviel Prozent das Alter die Kosten erklärt, wenn die Kosten normal verteilt sind.

Wir sehen die Regressionsgerade und die einzelnen Kostenpunkte. Wenn die einzelnen Kostenpunkte auf der Regressionsgerade liegen würden, wären die Kosten zu 100% durch das Alter erklärt.

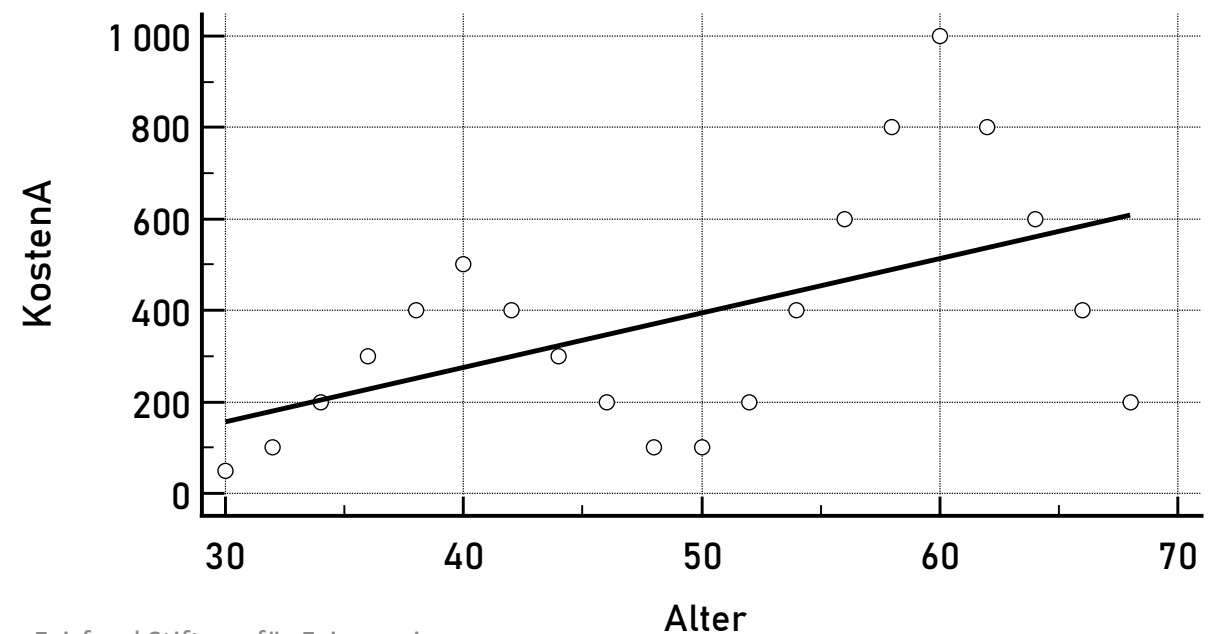
In diesem Beispiel ist dies offensichtlich nicht der Fall. Das Statistikprogramm gibt hier ein $r=0.53$ aus, das R^2 (man sagt R Quadrat) beträgt somit $0.53 \times 0.53 = 0.28$. In diesem fiktiven Beispiel erklärt das Alter somit 28% der Kosten oder *72% der Kosten NICHT (siehe später)*.

Erklärung der Kosten durch das Alter

$$y = -200.526 + 11.898 x$$

$n = 20$

$r = 0.53; P = 0.016$



Statistische Grundlagen für Juristinnen und Juristen zu Fragen der Wirtschaftlichkeitsverfahren in der Schweiz

Wirtschaftlichkeitsprüfungen und statistische Methoden

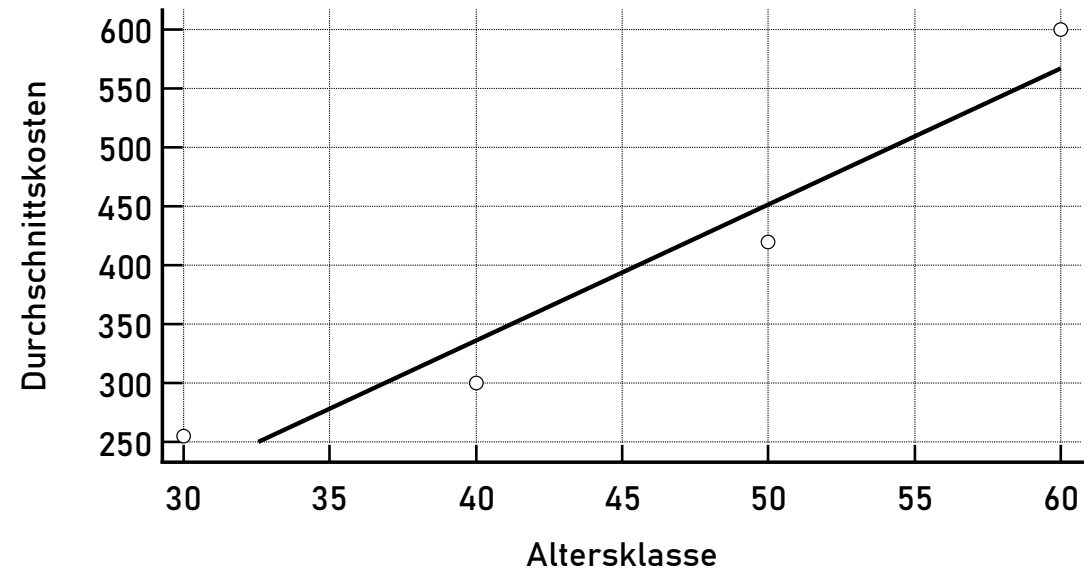
1.4.4. Beispiel zu Alter und Kosten

Jetzt wollen wir wissen, zu wieviel Prozent die Altersklassen die Durchschnittskosten erklären.

Wir sehen die Regressionsgerade und die 4 Kostenpunkte. Wenn die einzelnen Kostenpunkte auf der Regressionsgerade liegen würden, wären die Kosten zu 100% durch das Alter erklärt.

In diesem Beispiel ist dies fast erfüllt! Das Statistikprogramm gibt hier ein $r=0.97$ aus, das R^2 beträgt somit $0.97 \times 0.97 = 0.94$. In diesem fiktiven Beispiel erklären die Altersklassen somit 94% der Kosten.

Altersklassen erklären Durchschnittskosten
 $y = -126.000 + 11.550 x$
 $n = 4$
 $r = 0.97; P = 0.033$



Statistische Grundlagen für Juristinnen und Juristen zu Fragen der Wirtschaftlichkeitsverfahren in der Schweiz

Wirtschaftlichkeitsprüfungen und statistische Methoden

1.4.5. Beispiel zu Alter und Kosten

Wir haben jetzt gesehen, dass Einzeldaten aggregiert zu Mittelwerten das R^2 (der Erklärungsgehalt des statistischen Modells) massiv ansteigen lassen, nämlich in unserem Beispiel der normal verteilten Kosten bei 20 Personen unterschiedlichen Alters von 28% auf 94%.

Wie ist das möglich?

Wenn man Mittelwerte bildet, wird jede einzelne Person in diesen Mittelwert gepfercht, das heisst, die Streubreite der Einzeldaten verschwindet aus dem Modell, das R^2 (ein Mass für die Summe der Abstände von der Regressionslinie) steigt massiv an, weil die Abstände abnehmen.

Wird das R^2 der Gruppenmittelwerte kommuniziert, hat das Modell plötzlich einen sehr hohen Erklärungsgehalt. Das ist einer der Statistik-Tricks von santésuisse. Sie bilden, wie aus den WZW-Statistiken bekannt, Alters- und Geschlechtsgruppen und berechnen hier das R^2 und behaupten, dass damit ihr Modell ein sehr hohes R^2 habe.

Statistische Grundlagen für Juristinnen und Juristen zu Fragen der Wirtschaftlichkeitsverfahren in der Schweiz

Wirtschaftlichkeitsprüfungen und statistische Methoden

1.4.6. Beispiel zu Alter und Kosten

Hier sehen wir eine Folie von santésuisse. Aufgelistet sind die Facharztgruppen, die Anzahl Ärzte (ZSR) und das R^2 . Dieses beträgt bei allg. Innere Medizin 82%. Es wurden Alters- und Geschlechtsklassen gebildet, um die Kosten zu erklären.

Diese Statistik sagt nichts aus, wie wir gesehen haben, da das R^2 aus aggregierten Mittelwerten besteht und nicht aus Einzeldaten. Dies ist ein Fall von Statistikmissbrauch.

Statistikdaten 2011									
FAG-Code	FAG	R Squared	Adjusted R Squ	Anzahl ZSR	FAG-Code	FAG	R Squared	Adjusted R S	Anzahl ZSR
101	Anästhesiologie	0.88	0.874	132	136	Hämatologie	0.938	0.934	45
102	Chirurgie	0.717	0.708	439	145	Gastroenterologie	0.49	0.472	203
103	Dermatologie und Vener	0.634	0.623	363	150	Radio-Onkologie u	0.88	0.852	10
104	Gynäkologie und Geburt	0.901	0.896	1'030	151	Nuklearmedizin	0.952	0.93	4
106	Endokrinologie und Diab	0.697	0.684	107	152	Physikalische Med	0.673	0.656	57
107	Pneumologie	0.569	0.553	146	153	Praktischer Arzt / A	0.787	0.781	935
108	Neurochirurgie	0.758	0.743	57	154	Allergologie und k	0.603	0.585	65
109	Neurologie	0.524	0.506	222	155	Arbeitsmedizin	0.846	0.768	4
110	Psychiatrie und Psychoth	0.442	0.417	2'025	156	Herz- und thorakal	0.809	0.759	10
111	Kinder- und Jugendpsych	0.416	0.379	275	157	Infektiologie	0.859	0.842	19
113	Ophthalmologie	0.857	0.853	669	158	Intensivmedizin	0.645	0.508	5
114	Orthopädische Chirurgie	0.664	0.654	476	159	Kiefer- und Gesich	0.495	0.465	50
115	Oto-Rhino-Laryngologie	0.696	0.687	311	160	Kinderchirurgie	0.831	0.783	16
116	Kinder- und Jugendmedi	0.751	0.734	875	161	Klinische Pharmak	1		1
117	Radiologie	0.908	0.905	162	163	Medizinische Onk	0.758	0.747	126
118	Tropen- und Reisemediz	0.811	0.792	15	164	Nephrologie	0.714	0.697	67
119	Urologie	0.717	0.706	159	165	Pathologie	0.896	0.888	23
120	Rheumatologie	0.516	0.5	307	166	Pharmazeutische M	1		1
121	Allgemein Innere Mediz	0.826	0.821	5'331	167	Plastische, Rekons	0.683	0.67	106
122	Angiologie	0.731	0.719	83	168	Prävention und Ge	0.882	0.728	2
135	Kardiologie	0.726	0.716	337	175	Gruppenpraxen	0.835	0.829	129

Folie 5

Statistische Grundlagen für Juristinnen und Juristen zu Fragen der Wirtschaftlichkeitsverfahren in der Schweiz

Wirtschaftlichkeitsprüfungen und statistische Methoden

1.5.1. Prüfung der Zuverlässigkeit des statistischen Regressionsmodells

Wir haben in 1.4.3. und 1.4.4. geschaut, wie gut das Modell (die Regressionsgleichung) die Kosten erklärt. Das Regressionsmodell lässt einen Modellwert berechnen. Die Formel aus 1.4.3 lautet: $\text{Kosten (y)} = -200.526 + 11.898 \times (\text{Alter})$. **Diese Formel hat das Statistik Programm erstellt, um die erwarteten Kosten (y) aus dem Alter (x) zu berechnen.**

Erinnern wir uns: Das Alter soll die Kosten erklären. Wir setzen jetzt also in der Formel für x das Alter ein, z.B. 30 Jahre. Wir wissen, dass das Resultat 50 lauten sollte (siehe 1.2.). Jetzt rechnen wir also $30 \times 11.898 = 356.940$. Davon müssen wir nun gemäss Formel 200.526 abziehen, ergibt 156.414.

Die Formel sollte also hier den Wert 50 ergeben, ergibt aber den Wert 156. Total daneben. Ab wann ist der Wert aber daneben? In der Literatur wird für das Physician Profiling eine Reliability von 90% gefordert. Das bedeutet, dass der Modellwert den Realwert $\pm 5\%$ abbilden sollte. Beträgt also der Realwert 50, so ist der Modellwert zuverlässig (reliable), wenn er zwischen 47.5 und 52.5 liegt. Der Modellwert betrug aber 156 und ist damit nicht zuverlässig. Bei den 20 Realwerten zeigen die Berechnungen, dass nur 1:20 im Reliability-Bereich liegt. Die Reliability beträgt somit 5%, das Modell ist unbrauchbar! Das Alter erklärt nicht die Kosten!

Statistische Grundlagen für Juristinnen und Juristen zu Fragen der Wirtschaftlichkeitsverfahren in der Schweiz

Wirtschaftlichkeitsprüfungen und statistische Methoden

1.5.2. Prüfung der Zuverlässigkeit des statistischen Regressionsmodells

In dieser Excel Tabelle wird für jedes Alter der Realwert der Kosten und der Modellwert der Kosten aufgrund der Regressionsformel berechnet.

Die mittlere Spalte ist die Abweichung des Modellwerts vom Realwert. Ferner wird die 90% Reliability für jeden Realwert berechnet.

Liegt der Modellwert ausserhalb des Reliability-Wertes, ist er «unreliable», das Ergebnis ist ungenau, obwohl das Programm richtig rechnet.

Reliability bei normaler Verteilung

Reliability bei normaler Verteilung (links)						Reliability bei normaler Verteilung (rechts)					
Reliability	Abweichung des Modellwerts vom Realwert					Reliability	Abweichung des Modellwerts vom Realwert				
KostenA	Formel	-200.526	11.898			KostenA	Formel	-126	11.55		
Realwert	Modellwert	Abweichung	90%Reliable		Reliability	Realwert	Modellwert	Abweichung	90%Reliable		Reliability
50	156	106	47.5	52.5	0	50	221	171	47.5	52.5	0
100	180	80	95	105	0	100	244	144	95	105	0
200	204	4	190	210	1	200	267	67	190	210	0
300	228	-72	285	315	0	300	290	-10	285	315	1
400	252	-148	380	420	0	400	313	-87	380	420	0
500	275	-225	475	525	0	500	336	-164	475	525	0
400	299	-101	380	420	0	400	359	-41	380	420	0
300	323	23	285	315	0	300	382	82	285	315	0
200	347	147	190	210	0	200	405	205	190	210	0
100	371	271	95	105	0	100	428	328	95	105	0
100	394	294	95	105	0	100	452	352	95	105	0
200	418	218	190	210	0	200	475	275	190	210	0
400	442	42	380	420	0	400	498	98	380	420	0
600	466	-134	570	630	0	600	521	-79	570	630	0
800	490	-310	760	840	0	800	544	-256	760	840	0
1000	513	-487	950	1050	0	1000	567	-433	950	1050	0
800	537	-263	760	840	0	800	590	-210	760	840	0
600	561	-39	570	630	0	600	613	13	570	630	1
400	585	185	380	420	0	400	636	236	380	420	0
200	609	409	190	210	0	200	659	459	190	210	0
Summe					1	Summe					2
Alle					20	Alle					20
Prozent Reliability					5	Prozent Reliability					10

Statistische Grundlagen für Juristinnen und Juristen zu Fragen der Wirtschaftlichkeitsverfahren in der Schweiz

Wirtschaftlichkeitsprüfungen und statistische Methoden

1.6. Beispiel zu Alter und Kosten, schiefe Verteilung

Wir wiederholen jetzt unseren Lehrgang, jedoch mit neuer Kostenverteilung pro Patienten.

Patienten Nr. und Alter bleiben also gleich wie in 1.2, jedoch steigen die Kosten als Heavy Tail Verteilung massiv an. Diese Kostenverteilung ist bei vielen Facharztgruppen anzutreffen: die meisten kosten nichts oder sehr wenig, wenige kosten sehr viel (Pareto oder Heavy-Tail Verteilung der Kosten).

Jetzt führen wir die gleichen Schritte durch wie mit den normal verteilten Kosten.

Patient Nr	Alter	Heavy Tail
1	30	0
2	32	0
3	34	0
4	36	0
5	38	0
6	40	5
7	42	6
8	44	7
9	46	8
10	48	10
11	50	12
12	52	14
13	54	16
14	56	18
15	58	25
16	60	50
17	62	100
18	64	300
19	66	1000
20	68	5000

Statistische Grundlagen für Juristinnen und Juristen zu Fragen der Wirtschaftlichkeitsverfahren in der Schweiz

Wirtschaftlichkeitsprüfungen und statistische Methoden

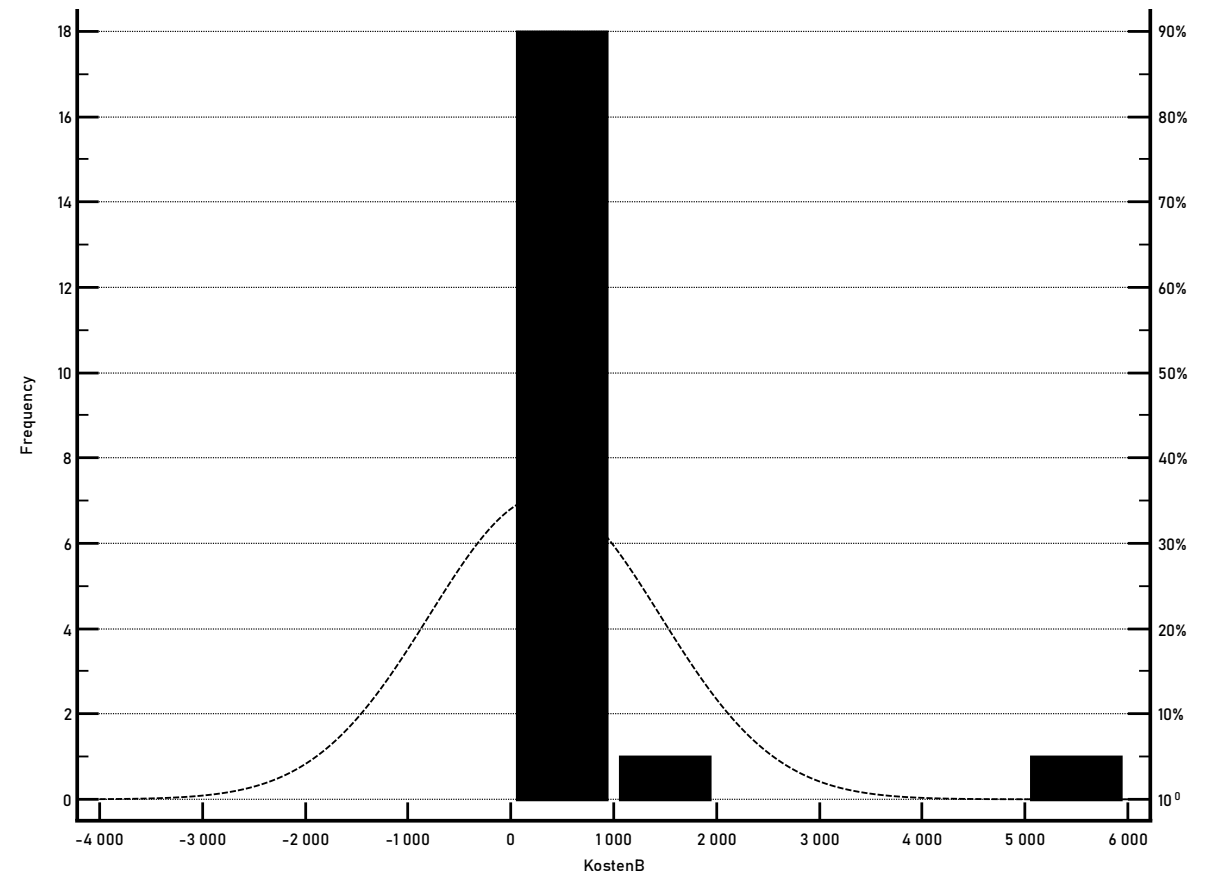
1.6.1. Beispiel zu Alter und Kosten, schiefe Verteilung

Die Analyse der Kostendaten (KostenB in der x-Achse) zeigt, dass die meisten Personen gar nicht kosten, wenige Kosten sehr viel.

Die Daten sind nicht normal verteilt, sondern Pareto verteilt und zwar hoch signifikant nicht normal verteilt ($p < 0.00001$).

Mit dieser Verteilung sind die Regressionsmodelle überfordert bzw. überhaupt nicht anwendbar.

In den folgenden Folien machen wir's trotzdem, ungeniert, wie santésuisse.



Statistische Grundlagen für Juristinnen und Juristen zu Fragen der Wirtschaftlichkeitsverfahren in der Schweiz

Wirtschaftlichkeitsprüfungen und statistische Methoden

1.6.2. Beispiel zu Alter und Kosten, schiefe Verteilung

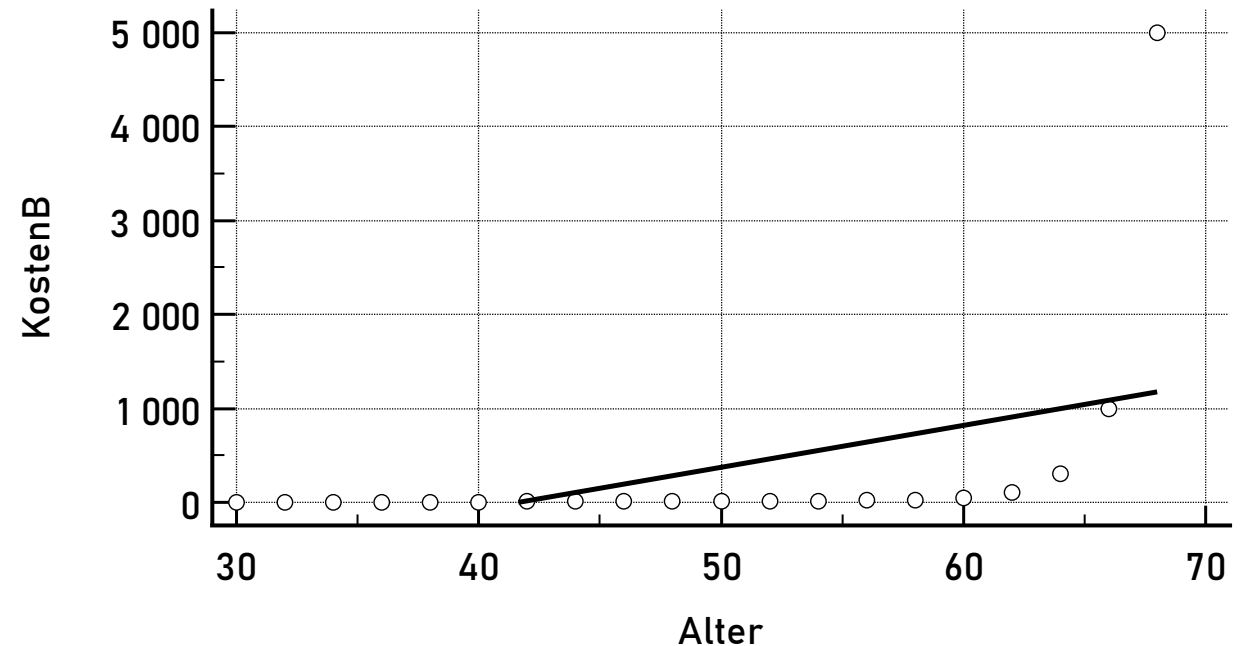
Wir sehen hier die Regressionsgerade für 20 Personen und die 20 Kosten (KostenB). Das R^2 beträgt hier 22%.

Die Regressionsgerade berechnet für jedes Alter die modellierten Kosten (das Kostenmodell von santésuisse!). Nehmen wir ein Beispiel und benutzen das Statistikprogramm, welches folgende Formel liefert:

$$\text{Kosten (y)} = -1857.642 + 44.616 x$$

Ergebnis für 68 Jahre = 1'221 statt 5'000, für 30 Jahre minus 474 statt 0. Mit dieser Fehlerbreite kämpft santésuisse.

Regression bei schief verteilten Kosten
 $y = -1857.642 + 44.616 x$
 $n = 20$
 $r = 0.47; P = 0.036$



Statistische Grundlagen für Juristinnen und Juristen zu Fragen der Wirtschaftlichkeitsverfahren in der Schweiz

Wirtschaftlichkeitsprüfungen und statistische Methoden

Zahlenreihe B (nicht normal verteilt)

Altersklasse	Alterswert	Mittelwert
30-38	30	0
40-48	40	7
50-58	50	17
60-68	60	1290

1.6.3. Beispiel zu Alter und Kosten, schiefe Verteilung

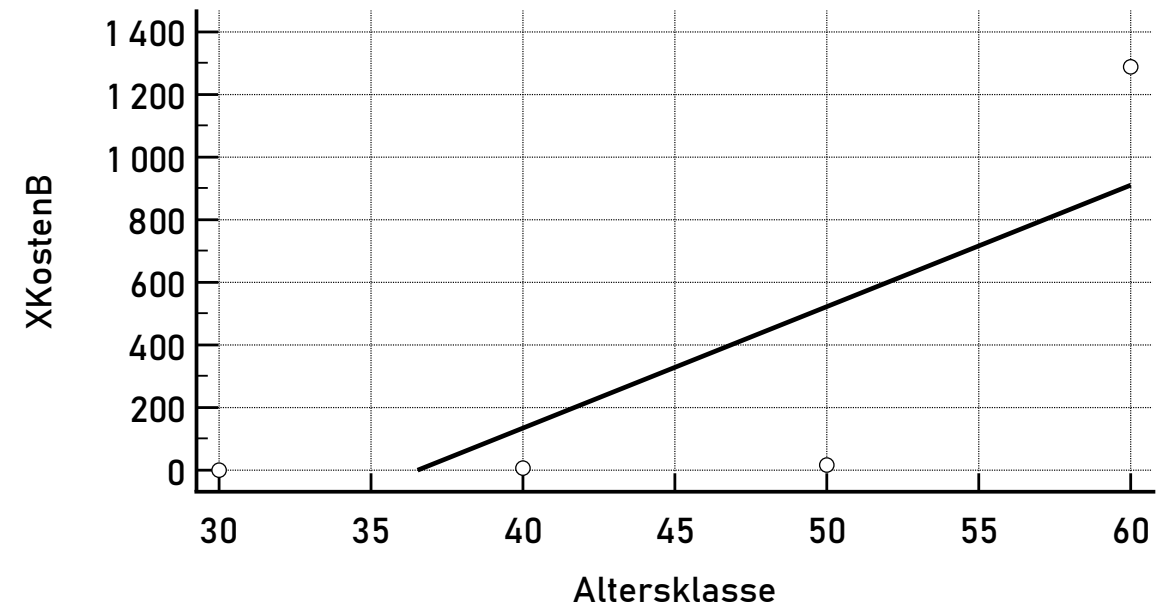
Wir sehen hier die Regressionsgerade für 4 Altersklassen und die 4 Kostenmittelwerte. Das R^2 beträgt hier 61%.

Die Regressionsformel der aggregierten Durchschnitts-Daten lautet:

$$\text{Kosten (y)} = -1417.5 + 38.8 x.$$

Wir multiplizieren also das Alter mit 38.8 und ziehen davon 1417.5 ab. Für 68 Jahre ergibt sich erneut 1'221, für 30 Jahre -253, für 60 Jahre 988 statt real 50 (Vergleiche 1.5.). Die Rückrechnung in die Realität ist somit massiv verfälscht. Dies wird später noch diskutiert.

Regression Altersklassen KostenB
 $y = -1417.500 + 38.800 x$
 $n = 4$
 $r = 0.78; P = 0.219$



Statistische Grundlagen für Juristinnen und Juristen zu Fragen der Wirtschaftlichkeitsverfahren in der Schweiz

Wirtschaftlichkeitsprüfungen und statistische Methoden

1.6.4. Reliability Test bei schiefer Kostenverteilung

Wir berechnen nun die Reliability gleich wie in 1.5.2. Das Ergebnis: Kein Modellwert lag richtig, die Reliability beträgt 0% und ist damit für schiefe verteilte Kosten noch schlechter als für normal verteilte Kosten.

Reliability bei nicht normaler Verteilung

Reliability bei nicht normaler Verteilung (links)					Reliability bei nicht normaler Verteilung (rechts)				
Reliability	Abweichung des Modellwerts vom Realwert				Reliability	Abweichung des Modellwerts vom Realwert			
KostenB	Formel	-1857.642	44.616		KostenB	Formel	-1417.5	38.8	
Realwert	Modellwert	Abweichung	90%Reliable	Reliability	Realwert	Modellwert	Abweichung	90%Reliable	Reliability
0	-519	-519	0	0	0	-254	-254	0	0
0	-430	-430	0	0	0	-176	-176	0	0
0	-341	-341	0	0	0	-98	-98	0	0
0	-251	-251	0	0	0	-21	-21	0	0
0	-162	-162	0	0	0	57	57	0	0
5	-73	-78	4.75	5.25	0	135	130	4.75	5.25
6	16	10	5.7	6.3	0	212	206	5.7	6.3
7	105	98	6.65	7.35	0	290	283	6.65	7.35
8	195	187	7.6	8.4	0	367	359	7.6	8.4
10	284	274	9.5	10.5	0	445	435	9.5	10.5
12	373	361	11.4	12.6	0	523	511	11.4	12.6
14	462	448	13.3	14.7	0	600	586	13.3	14.7
16	552	536	15.2	16.8	0	678	662	15.2	16.8
18	641	623	17.1	18.9	0	755	737	17.1	18.9
25	730	705	23.75	26.25	0	833	808	23.75	26.25
50	819	769	47.5	52.5	0	911	861	47.5	52.5
100	909	809	95	105	0	1000	988	95	105
300	998	698	285	315	0	3000	1066	766	285
1000	1087	87	950	1050	0	10000	1143	143	950
5000	1176	-3824	4750	5250	0	50000	1221	-3779	4750
Summe				0	Summe				0
Alle				20	Alle				20
Prozent Reliability				0	Prozent Reliability				0

Statistische Grundlagen für Juristinnen und Juristen zu Fragen der Wirtschaftlichkeitsverfahren in der Schweiz

Wirtschaftlichkeitsprüfungen und statistische Methoden

1.6.5. Winsorisierung

Die Winsorisierung ist ein bei solchen Analysen standardmässig durchgeführter Trick, um das Problem der schiefen Verteilung zu umgehen. Er wird in der internationalen Literatur praktisch immer verwendet, um extrem hohe Kosten einer erkrankten Person zu entfernen.

In unserem Beispiel würden sinnvollerweise der 68-jährigen Patienten mit den 5'000 CHF Kosten aus der Statistik entfernen, je nach Modell auch der 66-jährigen mit den 1'000 CHF.

Die FMH hat mit santésuisse vereinbart, diesen Standard, d.h. die Winsorisierung, nicht anzuwenden.

Statistische Grundlagen für Juristinnen und Juristen zu Fragen der Wirtschaftlichkeitsverfahren in der Schweiz

Wirtschaftlichkeitsprüfungen und statistische Methoden

1.7. Regressionsmodell besser verstehen

Anhand von 20 Alters- und Kostendaten, einmal mit normaler, einmal mit schiefer Verteilung der Kosten, haben wir verstanden, wie die statistische Maschine von santésuisse funktioniert und wie mit der Aggregation von Einzeldaten zu Gruppen (z.B. Altersklassen) der Erklärungsgehalt des verwendeten Modells künstlich erhöht wird und damit ein nur scheinbar hoher Erklärungsgehalt erzeugt wird, um die Glaubwürdigkeit der Methode zu erhöhen. Basierend auf diesem Wissen, können nun die folgenden Ausführungen besser verstanden werden.

Statistische Grundlagen für Juristinnen und Juristen zu Fragen der Wirtschaftlichkeitsverfahren in der Schweiz

Wirtschaftlichkeitsprüfungen und statistische Methoden

1.7.1. Regressionsmodell besser verstehen

Der RSS-Index

Der RSS-Index ist eine reine Kostendurchschnittsmethode. Hier wird kein ANOVA- oder Regressionsmodell verwendet, um die Kosten zu erklären. Die Statistik besteht aus einem einfachen Durchschnittswert. Wird dieser überschritten (>130%), spricht santésuisse von Unwirtschaftlichkeit.

Die Reliability von 130% ist unbekannt. Der Wert dürfte zuverlässiger sein, wenn die behandelten Krankheiten in etwa gleich sind, so will es der Gesetzgeber. Dass diese Vergleichbarkeit auf der Facharztebene (FAG) inexistent ist, werden wir später anhand der DDD / PCG (pharmazeutischen Kostengruppen) herleiten.

Da die Durchschnittskosten-Methode lediglich die Kosten analysiert, prüft sie nicht die Wirtschaftlichkeit. Der RSS-Index prüft nicht die Wirtschaftlichkeit, sondern überdurchschnittliche Kosten, welche dann als unwirtschaftlich betrachtet werden. Ein Fall von Etikettenschwindel.

Statistische Grundlagen für Juristinnen und Juristen zu Fragen der Wirtschaftlichkeitsverfahren in der Schweiz

Wirtschaftlichkeitsprüfungen und statistische Methoden

1.7.2. Regressionsmodell besser verstehen

Der ANOVA-Index

Der ANOVA-Index ist eine «Weiterentwicklung» des RSS-Indexes, indem erstmals für die Schweiz bei den Wirtschaftlichkeitsverfahren «erklärende», also unabhängige Variablen, den RSS-Index «korrigieren».

Im Wesentlichen wurde anhand von Alters- und Geschlechtsgruppen der RSS-Index korrigiert. Wir haben die Daten von 5'224 einzelnen internistischen Arztpraxen aus dem Jahr 2016 und schauen uns an, wie das Alter die direkten Arztkosten erklärt.

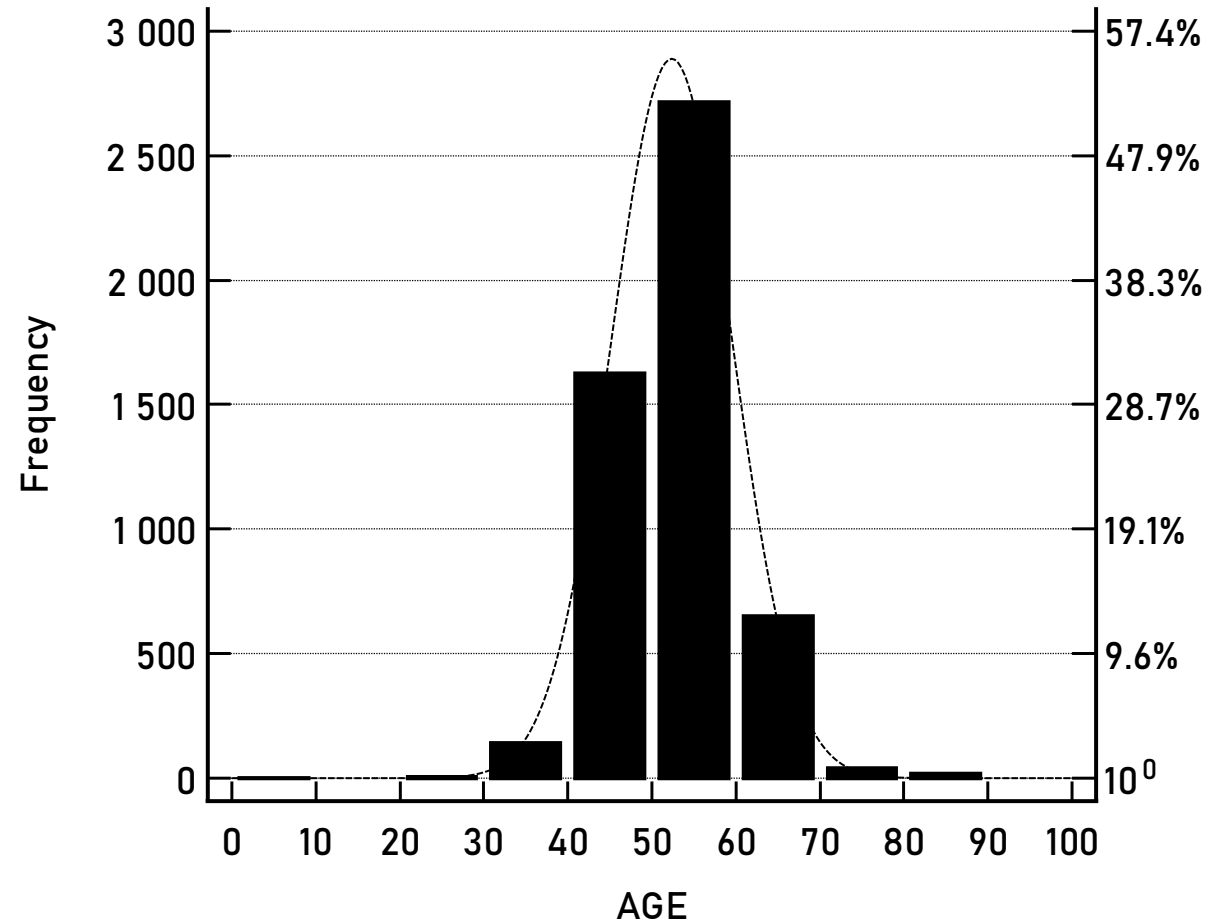
Statistische Grundlagen für Juristinnen und Juristen zu Fragen der Wirtschaftlichkeitsverfahren in der Schweiz

Wirtschaftlichkeitsprüfungen und statistische Methoden

1.7.3. Regressionsmodell besser verstehen

Durchschnittsalter pro Arztpraxis

Die Analyse mit MedCalc® zeigt die Verteilung des Durchschnittsalters pro Arztpraxis (5'225 Arztpraxen, minimal 6 Jahre, maximal 88 Jahre, Durchschnitt 52 Jahre), 4'892'944 Patienten und die Verteilung ist schief, also nicht normal verteilt ($p < 0.0001$)!



Statistische Grundlagen für Juristinnen und Juristen zu Fragen der Wirtschaftlichkeitsverfahren in der Schweiz

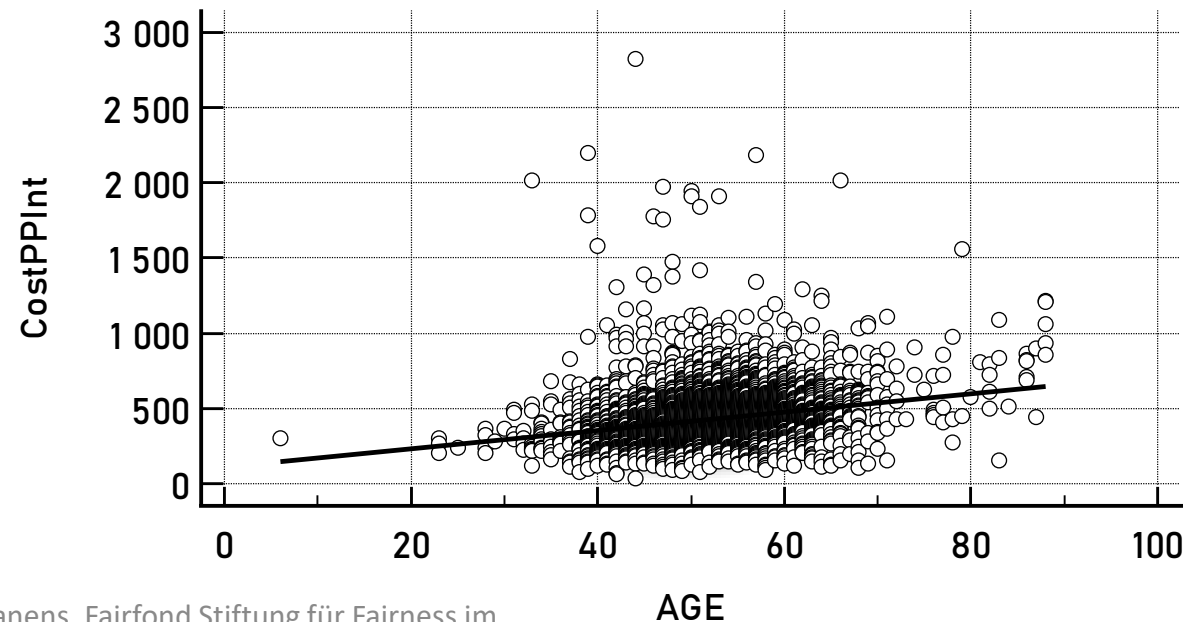
Wirtschaftlichkeitsprüfungen und statistische Methoden

1.7.3.1 Regressionsmodell besser verstehen

Durchschnittsalter pro Arztpraxis

Die Analyse mit MedCalc® zeigt, dass das Durchschnittsalter 6% der direkten Kosten/Patient erklärt.

Direkte Kosten und mittleres Alter pro Praxis
 $y = 109.847 + 6.096 x$
 $n = 5224$
 $r = 0.25; P < 0.001$



Statistische Grundlagen für Juristinnen und Juristen zu Fragen der Wirtschaftlichkeitsverfahren in der Schweiz

Wirtschaftlichkeitsprüfungen und statistische Methoden

1.7.3.2. Regressionsmodell besser verstehen

Tabelle 1.7.3.2. Sie sehen nur einen Ausschnitt von 5'224 Datensätzen

All 5224
 Reliable 670
 % Reliable 12.8

Reliability Test

Der vom Regressionsmodell erzeugte Modellwert zur Erklärung der Kosten aufgrund des Durchschnittsalters von 5'224 Praxen ist in 87% der Fälle nicht zuverlässig. Nur bei 13% der Arztpraxen stimmt die Formel mit der Realität überein (Datenquelle: santésuisse).

Im ANOVA Index wurde der RSS-Index für die Kosten von Alters- und Geschlechtsgruppen «korrigiert». Aufgrund dieser Analyse ist diese Korrektur unzuverlässig, das heisst, es erhalten Arztpraxen eine Reduktion des RSS-Index, der nicht stattfinden sollte, und umgekehrt.

Alter	Realwert	Modellwert	Abweichung des Modellwerts vom Realwert		90%Reliable	Reliability
			KostenB	Formel		
			109.847	6.096		
50	323	415	92	306.85	339.15	0
45	198	384	186	188.1	207.9	0
60	565	476	-89	536.75	593.25	0
39	546	348	-198	518.7	573.3	0
35	214	323	109	203.3	224.7	0
53	389	433	44	369.55	408.45	0
55	537	445	-92	510.15	563.85	0
45	360	384	24	342	378	0
39	402	348	-54	381.9	422.1	0
55	397	445	48	377.15	416.85	0
51	569	421	-148	540.55	597.45	0
61	653	482	-171	620.35	685.65	0
65	522	506	-16	495.9	548.1	1
46	749	390	-359	711.55	786.45	0
53	437	433	-4	415.15	458.85	1

Statistische Grundlagen für Juristinnen und Juristen zu Fragen der Wirtschaftlichkeitsverfahren in der Schweiz

Wirtschaftlichkeitsprüfungen und statistische Methoden

1.7.3.4. Regressionsmodell besser verstehen

Reliability Test mit multiplem Regressionsmodell

Wir haben gerade gesehen, dass das Alter die Kosten zuverlässig in 13% erklärt. Wählen wir nun mehrere Variablen aus der santésuisse Datenbank aus, um die Kosten zu erklären: verfügbar sind im Modell AGE, LAB, PHYSIO, MEDICOST, also das Durchschnittsalter pro Arztpraxis und die Kosten pro Arztpraxis für Labor, Physiotherapie und alle Medikamente.

Wir erwarten, dass die Zahl zuverlässig erklärter Arztkosten (interne Kosten pro Arztpraxis, ohne veranlasste Kosten) mit diesen Zusatzinformationen ansteigt. Wir haben wiederum mit MedCalc® ein Modell erzeugt, das Ergebnis lautet:

Constant	AGE	LAB	PHYSIO	MEDICOST
131.5316	4.419	0.5799	-0.1682	0.07978

Statistische Grundlagen für Juristinnen und Juristen zu Fragen der Wirtschaftlichkeitsverfahren in der Schweiz

Wirtschaftlichkeitsprüfungen und statistische Methoden

1.7.4.5. Regressionsmodell besser verstehen

Reliability Test mit multiplem Regressionsmodell: Ergebnis

Die Formel lautet also:

$$\text{COSTPPINT (y)} = 131.5316 \text{ (KONSTANTE)} + (\text{AGE} \times 4.419) + (\text{LAB} \times 0.5799) + (\text{PHYSIO} \times -0.1682) + (\text{MEDICOSTPP} \times 0.07978).$$

Die Summe der realen Kosten betrug 2'241'451, die Summe aus den Modellkosten ergab 2'241'443, das Modell hat also einen Fehler von 8 Franken auf 2'241'451 Mio Franken, also minimaler Fehler betreffend die Gesamtkosten.

Die Reliability auf der Praxisebene betrug weiterhin nur statt 12.8% neu 14.7%, womit sich die Information aus den zusätzlichen Morbiditätsvariablen LAB, PHYSIO, MEDICOSTPP um lediglich 1.9% verbesserte. Das R^2 (der Erklärungsgehalt des Modells) beträgt für dieses Modell nur 18% und entspricht für diesen Variablensatz den Erwartungen aus der Literatur! Im Ergebnis finden wir also eine «Verbesserung mit ANOVA von ganz schlecht (RSS-Index) auf schlecht.

Statistische Grundlagen für Juristinnen und Juristen zu Fragen der Wirtschaftlichkeitsverfahren in der Schweiz

Wirtschaftlichkeitsprüfungen und statistische Methoden

1.7.3.6. Regressionsmodell besser verstehen

Reliability Test mit multiplem Regressionsmodell: Ergebnis

Im Ergebnis finden wir also eine «Verbesserung mit ANOVA von ganz schlecht (RSS-Index) auf schlecht (ANOVA-Index). Solche Ergebnisse als Verbesserung zu verkaufen ist irreführend.

Statistische Grundlagen für Juristinnen und Juristen zu Fragen der Wirtschaftlichkeitsverfahren in der Schweiz

Wirtschaftlichkeitsprüfungen und statistische Methoden

1.7.4. Regressionsmodell besser verstehen

Der Regressionsindex

Wir haben also gesehen, dass der Erklärungsgehalt von RSS-Index und seine Verbesserung mit ANOVA-Index eine Verbesserung von «grottschlecht» zu immer noch ganz «schlecht» bewirkt, indem die Zahl der zuverlässigen Ergebnisse des Modells pro Arztpraxis von 13% auf 15% ansteigt und das R^2 mit ANOVA-Index 18% beträgt (vereinbartes Modell gemäss WZW-Vertrag vom 27.12.2013).

Diese Modelle erklären also nicht wirklich die Kosten, wie behauptet, die Berechnungen liegen praktisch nie punktgenau im Zentrum der Zielscheibe, sondern – teils massiv – daneben. Das statistische Modell gibt vor, etwas berechnen zu können, dass es nicht kann.

Im Folgenden schauen wir uns die «Verbesserung» im Regressionsindex bei 1'089 praktischen Ärzten aus dem Jahr 2019 an (alle Daten stammen von santésuisse).

Statistische Grundlagen für Juristinnen und Juristen zu Fragen der Wirtschaftlichkeitsverfahren in der Schweiz

Wirtschaftlichkeitsprüfungen und statistische Methoden

1.7.4.1. Regressionsmodell besser verstehen

Der Regressionsindex, multivariates lineares Regressionsmodell

Der ANOVA-Index wurde im Regressionsmodell durch weitere Variablen erweitert, welche die Kosten der Arztpraxen erklären sollen, als da sind: Hospitalisation im Vorjahr, Franchisenhöhe, Pharmazeutika (als DDD, bzw. pharmazeutische Kostengruppen).

Die vollständige Excel Tabelle zur FAG56 (praktischer Arzt) für das Jahr 2019 wurde ausgewertet. Es sind 1'089 Arztpraxen darin enthalten. Der Erklärungsgehalt der einzelnen Variablen auf die totalen Arztkosten pro Patienten (COSTPP) betrug für das Alter 7%, für die Franchise 10%, für die Hospitalisation im Vorjahr 8% und für DDD 42% (im Modell von santésuisse sind die gesamten Medikamentenkosten nicht enthalten, das R^2 beträgt für diese 67% und erklären so am meisten!).

Betreffend Reliability war das Alter mit 12.8% am besten, die anderen Variablen bewegten sich zwischen 0.6% bis 2.0%.

Statistische Grundlagen für Juristinnen und Juristen zu Fragen der Wirtschaftlichkeitsverfahren in der Schweiz

Wirtschaftlichkeitsprüfungen und statistische Methoden

1.7.4.2. Regressionsmodell besser verstehen

Der Regressionsindex, multivariates lineares Regressionsmodell

Für das gesamte Modell (Kosten $y = \text{Alter} + \text{Franchise} + \text{Hospitalisation} + \text{DDD}$) betrug das R^2 45% und die Reliability 11.6%. Die «Verbesserung» gegenüber dem ANOVA Modell betreffend R^2 beträgt somit $45\% - 18\% = 27\%$, die Reliability hat sich aber von 12.8% auf 11.6% verschlechtert. Somit werden die Kosten etwas besser erklärt, die Zuverlässigkeit bleibt ungenügend. Das Modell benötigt mehr Informationen, um den geforderten Genauigkeits-Standard zu erreichen (z.B. Morbiditätsvariablen, Diagnosedaten ecc)

Grund dafür ist der Irrtum im Modell, wonach die Gesamtkosten pro Patienten mit DDD erklärt werden können. Denn dies bedeutet automatisch, dass alle Medikamente, die nicht auf der PCG-Liste des Bundes auftauchen (2/3 aller Brands!!!) **nicht wirtschaftlich** sind. Das ist weder theoretisch noch empirisch nachvollziehbar. Ersetzen wir nun im multivariaten Regressionsmodell die DDD durch die Gesamtkosten für Medikamente (MEDIPP), stellen wir folgendes fest:

Für das gesamte Modell (Kosten $y = \text{Alter} + \text{Franchise} + \text{Hospitalisation} + \text{MEDIPP}$) betrug das R^2 69% statt 45% und die Reliability stieg von 11.6% auf 18.4% (+ 6.8%) an. Damit ist der Beweis erbracht, dass DDD ein technisches Problem im Regressionsmodell von santésuisse darstellen. Sie verschlechtern auch wie erwartet die Reliability.

Statistische Grundlagen für Juristinnen und Juristen zu Fragen der Wirtschaftlichkeitsverfahren in der Schweiz

Wirtschaftlichkeitsprüfungen und statistische Methoden

1.7.4.3. Regressionsmodell besser verstehen

Der Regressionsindex, Zusammenfassung

Wir haben mit einem einfachen Regressionsmodell nachgewiesen, was Polynomics in Ihrem Gutachten mit «komplexen statistischen Methoden» für die FAG Allg. Innere Medizin (Siehe Tabelle 31) herausgefunden hat.

Das Modell M1 entspricht etwa dem ANOVA Index, das Modell M3 dem Regressionsindex. Wir fanden für M1 ein R^2 von 18%, Polynomics fand 24%, für M3 fanden wir 45% und Polynomics fand 39%. Damit ist bewiesen, dass wir mit unserem einfachen Regressionsmodell die Berechnungen von Polynomics / santésuisse nachbilden können.

Tabelle 31 Erklärungsgehalt untersuchter Modelle mit Individualdaten

	Allgemeine Innere	Chirurgie	Gynäkologie	Kardiologie	Kinder/Jugend	Ophthalmologie
N	266'203	11'640	70'106	17'457	61'433	72'000
Adj. R2: Korrigiertes Bestimmtheitsmass						
▪ M1	0.24	0.16	0.10	0.14	0.12	0.31
▪ M2	0.34	0.16	0.12	0.16	0.15	0.39
▪ M3	0.39	0.20	0.13	0.24	0.17	0.40

Statistische Grundlagen für Juristinnen und Juristen zu Fragen der Wirtschaftlichkeitsverfahren in der Schweiz

Wirtschaftlichkeitsprüfungen und statistische Methoden

1.7.4.4. Regressionsmodell besser verstehen

Der Regressionsindex, Zusammenfassung

Wir haben aufgezeigt, wie die Aggregation von Daten (Mittelwert) durch die Entfernung der Variabilität zu einer künstlichen Erzeugung von sehr hohen R^2 Werten führt, wie dies Polynomics in Ihrem Bericht für die FAG Allg. Innere Medizin sehr schön belegen: R^2 88% statt 39% für Individualdaten.

Hier wird der Eindruck erweckt, das Regressionsmodell von Santésuisse hätte einen sehr hohen Erklärungsgehalt. Das ist Irreführung mit Statistik.

Tabelle 12 Erklärungsgehalt des Modells mit hohen Franchisen, Spital-im-Vorjahr und PCG

	Allgemeine Innere	Chirurgie	Gynäkologie	Kardiologie	Kinder/Jugend	Ophthalmologie	Psychiatrie/ Psychotherapie
N	205'430	15'543	23'401	14'295	17'115	34'238	59'562
N Ärzte	5'455	481	1'221	425	1'020	871	2'487
Adj. R^2							
Volles Modell	0.877	0.742	0.899	0.766	0.768	0.865	0.526
PE + AGG	0.817	0.729	0.892	0.700	0.745	0.862	0.427
PE	0.284	0.704	0.633	0.576	0.339	0.580	0.333

Statistische Grundlagen für Juristinnen und Juristen zu Fragen der Wirtschaftlichkeitsverfahren in der Schweiz

Wirtschaftlichkeitsprüfungen und statistische Methoden

1.7.4.5. Regressionsmodell besser verstehen

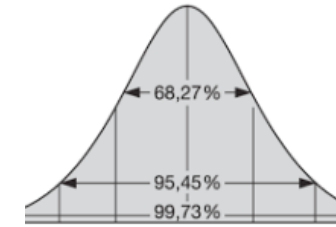
Erklärende Variablen versagen

Der Gesetzgeber hat zugelassen, überdurchschnittliche Kosten (im RSS-Index) automatisch als unwirtschaftlich zu bewerten, es sei denn, die Arztpraxis könne die hohen Kosten erklären.

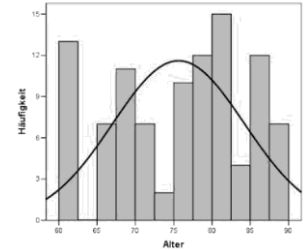
Auf Druck von FMH und Politik war santésuisse gezwungen, den RSS-Index zu verbessern und führte «erklärende» Variablen ein (ANOVA-Index, REGRESSIONS-Index). Wie wir gesehen haben, konnte damit eine Verbesserung von «grottschlecht» zu «ganz schlecht» erreicht werden. Es handelt sich um eine Scheinübung, wobei mit den DDD-Arztpraxen belohnt werden, die PCG-Medikamente verschreiben, ein Konstruktionsfehler. Oder sind nicht PCG gelistete Medikamente unwirtschaftlich? Zudem sind alle Daten von santésuisse nicht normal verteilt. Das Ergebnis ist entsprechend katastrophal. **Anmerkung:** Normal verteilt sind Daten, wenn die Messwerte innerhalb von 2 Standardabweichungen (95.45%) vom Mittelwert betragen (siehe Graphiken oben rechts).

Das Ergebnis all der Statistik ist somit: Der RSS-Index wird korrigiert, aber manchmal sogar in die falsche Richtung (DDD). Diese Irreführung könnte also juristische Konsequenzen haben. Gemäss der jüngsten Rechtsprechung genügt das Screening für die Bewertung als unwirtschaftlich nicht (BGE 150 V 129)

Normale Verteilung



Abnormale Verteilung



Statistische Grundlagen für Juristinnen und Juristen zu Fragen der Wirtschaftlichkeitsverfahren in der Schweiz

Wirtschaftlichkeitsprüfungen und statistische Methoden

2. Feststellung von Homogenität und Inhomogenität in Vergleichskollektiven

Die Statistik kennt verschiedene Verfahren und Formeln, um grosse Datenmengen auf das Homogenitätsproblem zu untersuchen.

Statistische Grundlagen für Juristinnen und Juristen zu Fragen der Wirtschaftlichkeitsverfahren in der Schweiz

Wirtschaftlichkeitsprüfungen und statistische Methoden

2.1. Mittelwert, Medianwert, Varianz, Standardabweichung

Definition des Mittelwertes: Summe der Einzelwerte dividiert durch Zahl der Werte

Beispiel: $150+100+50 = 3$ Einzelwerte = $300/3 = 100$

Statistische Beschreibung:

Stichprobengrösse	=	3	
Niedrigster Wert	=	50	
Höchster Wert	=	150	
Mittelwert	=	100	
Medianwert	=	100	Wert, der sich in der Mitte der Zahlenreihe befindet
Varianz	=	2500	Höhere Varianz bedeutet, dass sich Arztpraxen stärker unterscheiden [1]
Standard Abweichung (SD)	=	50	Wurzel aus der Varianz
Relative SD	=	50%	Mittelwert = 100%
Verteilung	=	normal	

[1] Berechnung der Varianz: Summe der Differenz der Einzelwerte vom Mittelwert im Quadrat = $-50^2 + 50^2 = 5000$ dividiert durch Anzahl Messungen minus 1 = $5000/2=2500$

Statistische Grundlagen für Juristinnen und Juristen zu Fragen der Wirtschaftlichkeitsverfahren in der Schweiz

Wirtschaftlichkeitsprüfungen und statistische Methoden

2.2. Mittelwert, Medianwert, Varianz, Standardabweichung

Selbstverständlich benutzen wir Computerprogramme wie MedCalc® oder Excel, um die Berechnung zu Standardabweichung und die Verteilung der Daten (normal verteilt versus nicht normal verteilt) zu beobachten.

Aufgrund der DDD in der FAG53 (praktischer Arzt) haben wir nun die Zahl Medikamente (DDD), welche pro PCG (pharmazeutische Kostengruppe) im Verlaufe des Indexjahres verordnet wurden, wobei die PCG direkte Rückschlüsse auf die Diagnosen erlauben (PCG-Insulin = Diabetes mellitus [1]).

Wir befragen nun unsere Statistikprogramme wie die einzelnen DDD in den Arztpraxen verteilt sind.

[1] 2/3 aller Medikamente sind nicht auf der PCG-Liste, so z.B. alle Blutverdünner (Thrombozyten-Aggregationshemmer, orale Antikoagulantien).

Statistische Grundlagen für Juristinnen und Juristen zu Fragen der Wirtschaftlichkeitsverfahren in der Schweiz

Wirtschaftlichkeitsprüfungen und statistische Methoden

2.3. Verteilung der DDD in der FAG53 (praktischer Arzt, 1'089 Arztpraxen, Jahr 2019)

KRK	Krebs komplex
MCR	Morbus Crohn / Colitis ulcerosa
MSK	Krankheiten des Gehirns oder des Rückenmarks: Multiple Sklerose
NIE	Nierenerkrankung
PAH	Pulmonale (arterielle) Hypertonie
PAR	Morbus Parkinson
PSO	Psoriasis
PSY	Psychose
RHE	Rheuma
SMC	Chronische Schmerzen (exklusive Opioiden)
SMN	Neuropathischer Schmerz
THY	Schilddrüsenerkrankungen
TRA	Transplantationen
WAS	Wachstumsstörung
ZFP	Zystische Fibrose / Pankreasenzyme
ZNS	Krankheiten des Gehirns oder des Rückenmarks: Sonstige

	KRK	MCR	MSK	NIE	PAH	PAR	PSO	PSY	RHE
Sample size	1089	1089	1089	1089	1089	1089	1089	1089	1089
Lowest value	0	0	0	0	0	0	0	0	0
Highest value	6.8	5.3	3.6	2.3	1.3	11.2	7.5	137.6	6.9
Arithmetic mean	0.06079	0.3339	0.02268	0.05014	0.01056	0.6503	0.3311	1.6482	0.4024
95% CI for the Arithmetic mean	0.04305 to 0.07853	0.3014 to 0.3664	0.01227 to 0.03309	0.04115 to 0.05912	0.005694 to 0.0154	0.5922 to 0.7084	0.2952 to 0.3670	1.3274 to 1.9691	0.3639 to 0.4409
Median	0	0.1	0	0	0	0.3	0.1	0.7	0.1
95% CI for the median	0.0000 to 0.0000	0.06676 to 0.1000	0.0000 to 0.0000	0.0000 to 0.0000	0.0000 to 0.0000	0.2000 to 0.3000	0.1000 to 0.1000	0.6000 to 0.7000	0.1000 to 0.2000
Variance	0.089	0.2982	0.03067	0.02283	0.006699	0.9553	0.3648	29.1197	0.4192
Standard deviation	0.2983	0.5461	0.1751	0.1511	0.08185	0.9774	0.604	5.3963	0.6474
Relative standard deviation	4.9076 (490.76%)	1.6356 (163.56%)	7.7213 (772.13%)	3.0138 (301.38%)	7.7506 (775.06%)	1.5029 (150.29%)	1.8240 (182.40%)	3.2740 (327.40%)	1.6090 (160.90%)
Standard error of the mean	0.00904	0.01655	0.005307	0.004579	0.00248	0.02962	0.0183	0.1635	0.01962
Coefficient of Skewness	14.6780 (P<0.0001)	2.8812 (P<0.0001)	14.5271 (P<0.0001)	6.5066 (P<0.0001)	10.7561 (P<0.0001)	3.5548 (P<0.0001)	4.5943 (P<0.0001)	17.0770 (P<0.0001)	3.6489 (P<0.0001)
Coefficient of Kurtosis	281.1163 (P<0.0001)	13.3291 (P<0.0001)	253.1946 (P<0.0001)	166.2955 (P<0.0001)	133.3871 (P<0.0001)	121.7492 (P<0.0001)	34.3001 (P<0.0001)	387.9693 (P<0.0001)	22.2131 (P<0.0001)
Shapiro-Wilk test	W=0.1805	W=0.6599	W=0.1079	W=0.3666	W=0.1109	W=0.6623	W=0.5652	W=0.2153	W=0.6403
for Normal distribution	reject Normality (P<0.0001)	reject Normality (P<0.0001)	reject Normality (P<0.0001)	reject Normality (P<0.0001)	reject Normality (P<0.0001)	reject Normality (P<0.0001)	reject Normality (P<0.0001)	reject Normality (P<0.0001)	reject Normality (P<0.0001)

	RHE	SMC	SMN	THY	TRA	WAS	ZFP	ZNS
Sample size	1089	1089	1089	1089	1089	1089	1089	1089
Lowest value	0	0	0	0	0	0	0	0
Highest value	6.9	51.1	16.4	32.6	4.7	0.5	4.3	63.9
Arithmetic mean	0.4024	9.6905	1.0283	5.7655	0.06465	0.002663	0.1785	1.1153
95% CI for the Arithmetic mean	0.3639 to 0.4409	9.2393 to 10.1418	0.9452 to 1.1113	5.5026 to 6.0283	0.04795 to 0.08134	0.001028 to 0.00429	0.1602 to 0.1969	0.9728 to 1.2578
Median	0.1	8.6	0.7	5.3	0	0	0.1	0.6
95% CI for the median	0.1000 to 0.2000	8.1000 to 9.1000	0.6000 to 0.7000	5.0000 to 5.6000	0.0000 to 0.0000	0.0000 to 0.0000	0.0000 to 0.1000	0.6000 to 0.7000
Variance	0.4192	57.5898	1.9508	19.5456	0.07887	0.0007558	0.09518	5.7439
Standard deviation	0.6474	7.5888	1.3967	4.421	0.2808	0.02749	0.3085	2.3966
Relative standard deviation	1.6090 (160.90%)	0.7831 (78.31%)	1.3583 (135.83%)	0.7668 (76.68%)	4.3442 (434.42%)	10.3234 (1032.34%)	1.7283 (172.83%)	2.1488 (214.88%)
Standard error of the mean	0.01962	0.23	0.04232	0.134	0.00851	0.0008331	0.009349	0.07263
Coefficient of Skewness	3.6489 (P<0.0001)	1.3686 (P<0.0001)	4.6154 (P<0.0001)	1.3386 (P<0.0001)	8.7433 (P<0.0001)	14.0228 (P<0.0001)	4.3101 (P<0.0001)	17.5565 (P<0.0001)
Coefficient of Kurtosis	22.2131 (P<0.0001)	3.2100 (P<0.0001)	37.0099 (P<0.0001)	3.7404 (P<0.0001)	102.5419 (P<0.0001)	124.3165 (P<0.0001)	36.4470 (P<0.0001)	437.9463 (P<0.0001)
Shapiro-Wilk test	W=0.6403	W=0.9069	W=0.6438	W=0.9133	W=0.2367	W=0.0717	W=0.6025	W=0.3167
for Normal distribution	reject Normality (P<0.0001)	reject Normality (P<0.0001)	reject Normality (P<0.0001)	reject Normality (P<0.0001)	reject Normality (P<0.0001)	reject Normality (P<0.0001)	reject Normality (P<0.0001)	reject Normality (P<0.0001)

Statistische Grundlagen für Juristinnen und Juristen zu Fragen der Wirtschaftlichkeitsverfahren in der Schweiz

Wirtschaftlichkeitsprüfungen und statistische Methoden

2.4. Verteilung der DDD in der FAG53 (praktischer Arzt, 1'089 Arztpraxen, Jahr 2019)

Die Tabelle 5.3 zeigt die massive Inhomogenität der behandelten Krankheiten in dieser Facharztgruppe, die ja dafür bekannt ist, dass praktische Ärztinnen und Ärzte ganz verschiedene Krankheiten behandeln.

Ein statistischer Vergleich dieser Arztpraxen untereinander verbietet sich automatisch, weil die behandelten Krankheiten nicht vergleichbar sind, die prozentuale Standardabweichung erreicht sehr hohe Werte.

Die Voraussetzung für die Berechnung des Regressions-Indexes, die Feststellung von Überarztung anhand des RSS-Indexes oder des ANOVA-Indexes verbietet sich automatisch wegen der Inhomogenität der behandelten Krankheiten.

Statistische Grundlagen für Juristinnen und Juristen zu Fragen der Wirtschaftlichkeitsverfahren in der Schweiz

Wirtschaftlichkeitsprüfungen und statistische Methoden

2.5. Verteilung der DDD in der FAG53 (praktischer Arzt, 1'089 Arztpraxen, Jahr 2019)

Im Verfahren wegen überhöhten Durchschnittskosten muss nun die angeschuldigte Arztpraxis «erklären», wo seine Praxisbesonderheiten gegenüber der Vergleichsgruppe liegen.

Es sollte an diesem Punkt eigentlich allen klar sein, dass diese «Erklärung» der Praxisbesonderheiten unmöglich objektiv sein kann, weil die Voraussetzung hierfür, nämlich die Vergleichbarkeit der behandelten Krankheiten nicht existiert.

Dieser fundamentale Irrtum in den Wirtschaftlichkeitsprüfungen konnte nun mit den DDD aufgedeckt werden. Die Feststellung von Praxisbesonderheiten ergeben nur einen Sinn, wenn die Vergleichsgruppe homogen ist. Sie ist es nicht, weder in der Facharztgruppe noch generell in der Medizin.

Statistische Grundlagen für Juristinnen und Juristen zu Fragen der Wirtschaftlichkeitsverfahren in der Schweiz

Wirtschaftlichkeitsprüfungen und statistische Methoden

3. Vom Risikoausgleich (risk equalisation)

Es ist kaum bekannt, dass die Regressionsanalyse zur Erklärung von Kosten auch beim Risikoausgleich unter den Krankenkassen wirkt, auch hier gelten die gleichen «erklärenden» Variablen wie die Demographie (Alter, Geschlecht, Wohnkanton) und die Morbiditätsvariablen (Hospitalisation im Vorjahr, Franchisenhöhe, DDD).

Die Kostenstruktur jeder einzelnen Krankenkasse wird in der gemeinsamen Einrichtung KVG ermittelt und anhand des linearen Regressionsmodells die Ausgleichzahlungen unter den Kassen ermittelt. Es geht um Millionen Beträge, die jährlich hin- und hergeschoben werden.

Doch der Risikoausgleich funktioniert nicht, die Zahl der Kassenwechsel steigt und steigt, die Jagd nach kostengünstigen Versicherten ist unverändert hoch, die Kosten pro Kassenwechsel 800 Fr zu 1 Mio Kassenwechsel, es geht also auch ins Geld.

Statistische Grundlagen für Juristinnen und Juristen zu Fragen der Wirtschaftlichkeitsverfahren in der Schweiz

Wirtschaftlichkeitsprüfungen und statistische Methoden

3.1. Vom Risikoausgleich (risk equalisation)

Der Bund hat den Effekt und die Güte dieses Risikoausgleichs durch Institute beurteilen lassen. Es wurde dazu im Januar 2024 ein [Bericht publiziert](#).

Die Autoren schreiben auf Seite 72: « «Allerdings werden dabei von den PCG auch 20% der Versicherten ohne Kosten in den Folgejahren fälschlicherweise als solche mit Kosten identifiziert, was einer hohen Fehlerrate entspricht». Auf Seite 99: «Wie in Anhang 9.11 näher erläutert, führt das verwendete Modell zu Verzerrungen und möglichen Falschaussagen». Auf Seite 173: «Die Schätzung der Modellparameter und der daraus abgeleiteten Ausgleichssätze ist statistisch ineffizient; Versicherte mit ausserordentlich hohen Nettoleistungen (Ausreisser) haben einen viel höheren Einfluss auf die Schätzung als Versicherte mit durchschnittlichen Nettoleistungen; die Aussagekraft des geschätzten Modells ist beschränkt. Sämtliche Inferenzresultate, wie allfällige Analysen zur Signifikanz der Einflussfaktoren und Prognoseintervalle sind (bei Verwendung konventioneller Berechnungsmethoden) ungültig. Ebenfalls ungültig sind viele der üblicherweise verwendeten Modellgütemasse (u.a. auch solche, die in diesem Bericht aufgeführt werden)».

Statistische Grundlagen für Juristinnen und Juristen zu Fragen der Wirtschaftlichkeitsverfahren in der Schweiz

Wirtschaftlichkeitsprüfungen und statistische Methoden

3.2. Vom Risikoausgleich (risk equalisation)

Die Autoren fordern auf Seite 174: «Ein Modell, bei welchem die Modellannahmen voraussichtlich weniger stark verletzt werden, wäre das Hurdle-Modell (z. B. Cragg, 1971), mit Annahme einer Gamma-Verteilung für die Nettoleistungen > 0 , und unter Verwendung einer robusten Schätzmethode. Während eine solche Anpassung ein stringentes statistisches Modell mit erhöhter Aussagekraft ermöglichen würde, hat sie zum Nachteil, dass die entsprechenden Anpassungen in der Verordnung und bei den Berechnungen voraussichtlich komplex sind.

Statistische Grundlagen für Juristinnen und Juristen zu Fragen der Wirtschaftlichkeitsverfahren in der Schweiz

Wirtschaftlichkeitsprüfungen und statistische Methoden

3.3. Vom Risikoausgleich (risk equalisation)

Das Winterthurer Institut für Gesundheitsökonomie, ZHAW Institut für Datenanalyse und Prozessdesign, ZHAW Universität Luzern mit den AutorInnen Reto Bürgin, Michael Stucki, Christina Vetsch-Tzogiou, Lukas Kauer, Andreas Kohler, Anna Drewek, Christoph Thommen, Marcel Dettling und Simon Wieser **kommen also zum gleichen Schluss wie wir**: die Regressionsanalyse ist für die Erfassung der erwarteten Kosten ein unzuverlässiges, häufig zu falschen Ergebnissen führendes Modell.

Genau dieses Modell dient santésuisse für die Beurteilung der Wirtschaftlichkeit von Arztpraxen. Da die Arztpraxen viel kleinere Patientenzahlen aufweisen als die Krankenversicherer mit mehreren Millionen Versicherten **wirken sich die Modellfehler noch massiver aus**. Vor dem Hintergrund dieser Informationen ist nicht verständlich, warum die FMH nicht schon lange die Notbremse gezogen hat.

Statistische Grundlagen für Juristinnen und Juristen zu Fragen der Wirtschaftlichkeitsverfahren in der Schweiz

Wirtschaftlichkeitsprüfungen und statistische Methoden

4. Vom Risikoausgleich (risk equalisation) bei der Einzelfallprüfung

Der Bund hat den Effekt und die Güte dieses Risikoausgleichs durch Institute beurteilen lassen. Es wurde dazu im Januar 2024 ein [Bericht publiziert](#). Das Ergebnis dieses Berichts ist für das Regressionsmodell für Risikoausgleich Gemeinsame Einrichtung KVG und Wirtschaftlichkeitsverfahren santésuisse vernichtend.

Genau dieses Modell verwenden nun aber einzelne Krankenversicherer für die Analyse einzelner Tarmed-Positionen von Leistungserbringern, z.B. [Blacklight Analytics](#). Anstelle der Branchendaten von santésuisse wird das Modell zudem nur über die Daten einer Krankenversicherung, z.B. Helsana, durchgeführt.

Die Reliability der Resultate wird dadurch nochmals erheblich vermindert. Beispiel Radiologie: Ob eine bestimmte Tarmed-Position für die Bildgebung des Körpers adäquat ist oder nicht, kann doch nicht davon abhängig gemacht werden, ob diese Person PCG-gelistete Medikamente einnimmt (dann wäre sie aus der Sicht von Blacklight Analytics kränker und die Verwendung einer bestimmten Tarmed-Position gerechtfertigter). Die Krankenversicherer haben mit der Statistik Methode Risikoausgleich eine Einnahmequelle geschaffen. Je unpräziser die Ergebnisse der Prüfung, desto mehr erscheinen die Leistungserbringer als «auffällig».

Statistische Grundlagen für Juristinnen und Juristen zu Fragen der Wirtschaftlichkeitsverfahren in der Schweiz

Wirtschaftlichkeitsprüfungen und statistische Methoden

5. Zusammenfassung

Das Controlling der Krankenversicherer im Gesundheitswesen wird immer komplexer und undurchsichtiger. Die verwendeten Statistiken sind weder Juristinnen noch Leistungserbringern in der Regel zugänglich. Frei von der Bürde staatlicher Kontrolle (z.B. agiert das BAG so, als würde santésuisse nicht der Aufsicht des BAG unterstellt sein) treiben die Krankenversicherer die Rückforderungen aufgrund von statistischen Resultaten immer mehr in die Höhe. Es ist ein lukratives Geschäftsmodell geworden.

Im Grundsatz werden in einem linearen Regressionsmodell Variablen gefüttert, welche hohe Kosten der Krankenversicherer erklären und die Versicherungsrisiken abdecken sollen. Die Täuschung von Justiz und Politik geschieht da, wo hohe Kosten automatisch das Etikett der «Unwirtschaftlichkeit» erhalten.

Die Ergebnisse sind in der Regel unbrauchbar. Das Regressionsmodell der Krankenversicherer funktioniert weder im Risikoausgleich untereinander noch in der Beurteilung der Wirtschaftlichkeit der Leistungserbringer. Wer anderes behauptet, nimmt bewusst eine Täuschung der Öffentlichkeit in Kauf.