

Zur Vergleichbarkeit statistischer Zahlen

Von Dr. O. Schenker, Bern

1. Die Notwendigkeit von Vergleichen in der Statistik.
 2. Bedeutung der Wahrscheinlichkeiten zu Vergleichszwecken.
 3. Über die Schwierigkeiten, Geburtenwahrscheinlichkeiten zu konstruieren.
 4. Vergleichen über die Heiratshäufigkeit nach dem gegenseitigen Alter, die Tuberkulosen- und Krebssterblichkeit.
 5. Vergleichen anhand der Alterskombinationen zusammenlebender Ehegatten in der Schweiz.
-

1. Zwei soziale Erscheinungen oder Zustände können bloss dann miteinander verglichen werden, wenn sie mit demselben Mass gemessen worden sind, oder wenn diese Erscheinungen, bzw. Zustände, voneinander abhängen; in letzterem Falle braucht kein gemeinsames Mass vorhanden zu sein. Man kann die Zahl der Geburten in verschiedenen Gebieten einander gegenüberstellen, wobei das gemeinsame Mass durch den Begriff der Geburt gegeben ist; man kann das Einkommen einer Familie mit der davon abhängigen Wohndichte in Verbindung bringen, obwohl sich hier das Einkommen nicht durch die Wohndichte ausdrücken lässt. Solche durch das Hilfsmittel der Statistik angestellte Vergleiche können einen mehr oder weniger hohen Grad von Objektivität besitzen, sofern anhand von richtigen Voraussetzungen richtige Schlüsse gezogen werden. Den Ausgangspunkt für alle statistischen Vergleichen bilden immer die absoluten Zahlen, welche manchen Zwecken der Sozialpolitik genügen; man wird aber dabei nicht stehen bleiben, besonders dann nicht, wenn es sich darum handelt, soziale Regelmässigkeiten und Abhängigkeiten zu erforschen; hier wird man Verhältniszahlen und Mittelwerte zu Rate ziehen müssen. Es kann vorkommen, dass eine wirkliche soziale Erscheinung einer fingierten gegenübergestellt wird; so kann man die Sterblichkeit eines bestimmten Berufes unter der Voraussetzung berechnen, dass die Sterblichkeitsraten der verschiedenen Alter mit denjenigen der allgemeinen Bevölkerung übereinstimmen, wie es Westergaard getan hat ¹⁾. Man stösst auf diesem Wege auf den Begriff der positiven bzw. negativen Abhängigkeit.

2. Unter den statistischen Verhältniszahlen zeichnen sich vor allem die Wahrscheinlichkeiten (entstanden durch Gegenüberstellung der wirklichen und möglichen Fälle) zufolge ihres wissenschaftlichen Wertes aus. Zur Illustration greifen wir die folgenden Zahlen aus der schweizerischen Statistik heraus:

¹⁾ Dr. Gottlieb Schnapper-Arndt, Sozialstatistik, Leipzig 1908, S. 221.

1 Altersklassen	Todesfälle ¹⁾ 1910 und 1911	Wohn- bevölkerung ²⁾ am 1. Jan. 1911	Todesfälle ¹⁾ 1920 und 1921	Wohn- bevölkerung ²⁾ am 1. Jan. 1921
0—4	27.329	409.991	16.839	334.415
5—14.	3.894	769.346	3.686	754.262
15—19.	2.757	356.425	2.688	386.821
20—29.	6.692	618.995	6.115	653.257
30—39.	7.351	546.838	6.447	543.588
40—49.	8.751	422.025	8.772	488.202
50—59.	11.708	301.533	12.775	362.998
60—69.	18.612	209.736	17.562	226.646
70—79.	19.855	100.178	20.462	107.450
80 und mehr. .	9.168	20.673	10.164	24.234
Total	116.117	3.755.740	105.510	3.881.873

¹⁾ Schweizerische statistische Mitteilungen, 1924, 1. Heft, S. 9; die Totgeborenen sind natürlich nicht mitgerechnet, denn wer nicht gelebt hat, kann auch nicht sterben.
²⁾ Schweizerische statistische Mitteilungen, 1925, 6. Heft, S. 49* und S. 50*.

Diese absoluten Zahlen können über den *Sterblichkeitsgrad* in den verschiedenen Altersklassen keine richtige Auskunft geben; die Abhängigkeit der Sterblichkeit vom Alter kommt erst ungetrübt zum Ausdruck, wenn die Zahl der wirklichen Todesfälle mit der Zahl der möglichen, d. h. mit der Wohnbevölkerung der betreffenden Altersklasse, gemessen wird. Bei Vergleichen kommt das gemeinsame Mass ohne weiteres zur Geltung, indem man stets dieselbe Zahl der möglichen Sterbefälle zugrunde legt, z. B. 1000 ¹⁾. Die Vorteile unseres dezimalen Zahlensystems gestatten dies ohne jede Rechnung. Wir lassen die aus den gegebenen absoluten Zahlen berechneten *Sterbenswahrscheinlichkeiten* folgen:

2 Altersklassen	0—4	5—14	15—19	20—29	30—39	40—49	50—59	60—69	70—79	80 u. mehr	Total
Auf 1000 mögliche Sterbefälle sind jährlich zu erwarten nach den Beobachtungen von:											
1910—1911.	33	3	4	5	7	10	19	42	90	181	15
1920—1921.	25	2	3	5	6	9	17	38	87	173	13

¹⁾ Die sogenannten Sterblichkeitsintensitäten besitzen noch einen höhern Grad von Vergleichbarkeit als die Sterbenswahrscheinlichkeiten, weil sie nicht bloss auf die Einheit der Bevölkerung, sondern auch auf die Einheit der Zeit berechnet werden, indem man die Zahl der Todesfälle durch die Zeit dividiert, welche von den unter dem Sterblichkeitsrisiko stehenden Personen durchlebt wird. Die Intensität der Sterblichkeit ist in der «Zeitschrift für schweizerische Statistik und Volkswirtschaft» zu verschiedenen Malen behandelt worden, so von Dr. Bohren: Die Intensität der Sterblichkeit, bestimmt auf Grund der zwei ersten schweizerischen Sterbetafeln, 1903, II, 60—68, und Prof. Dr. Ch. Moser: Die Intensität der Sterblichkeit und die Intensitätsfunktion, 1906, II, 109—114.

3. Nicht alle Wahrscheinlichkeiten sind so leicht zu bilden wie die Sterbenswahrscheinlichkeit. Fragen wir z. B. nach der Wahrscheinlichkeit, dass eine Frau gegebenen Alters im Laufe eines Jahres ein Kind gebiert (Geburtenwahrscheinlichkeit)? Schon der Begriff der Geburt ist umstritten. Nach Würzburger ist die Geburt der Augenblick, in dem das Geschlecht der Leibesfrucht sichtbar zum Ausdruck kommt¹⁾. Wie soll man hier die Zahl der möglichen Fälle ermitteln und der Forderung gleich möglicher Fälle genügen? Die zweite Bedingung verlangt Frauen gleichen Alters und Zivilstandes, aber auch gleicher Stillfähigkeit; ferner muss man Aufschluss haben über die Geburten, welche vor dem sechsten Schwangerschaftsmonat stattgefunden haben sowie über das Datum der letzten Geburt. Bei Erstgeburten hat man ferner mit vorehelicher Empfängnis zu rechnen, welche die Geburtenwahrscheinlichkeit höher erscheinen lässt als sie wirklich ist; sodann hat man den Wechsel im Zivilstand während der Beobachtungszeit sowie den vorkommenden Todesfällen Aufmerksamkeit zu schenken, abgesehen von den Wanderungen; kurz, wir haben bei der Frage nach der Geburtenwahrscheinlichkeit einer ganzen Menge von Faktoren Rechnung zu tragen, welche von der Statistik nicht erfasst werden. Das Problem kann beim gegenwärtigen Stande unseres Wissens zu keiner befriedigenden Lösung führen. In der Schweiz werden eingehende Beobachtungen über die Geburtenhäufigkeit bloss vom statistischen Amte der Stadt Zürich gemacht. Wir haben auf Grund der entsprechenden Zahlen, welche das ausgezeichnet redigierte Jahrbuch dieses Amtes enthält, und unter verschiedenen Annahmen, besonders über die Dauer der Stillfähigkeit, Geburtenwahrscheinlichkeiten berechnet, können aber auf ihre Wiedergabe verzichten, da sie, besonders für die jüngern Alter, in hohem Grade von der Stillfähigkeit abhängen und hierzu die statistische Unterlage fehlt; zudem wissen wir nichts über den Umfang der nicht anzeigepflichtigen Geburten, also mit einer Schwangerschaftsdauer unter sechs Monaten. Man weiss, dass die Dauer der Stillfähigkeit, besonders in den Städten, im Abnehmen begriffen ist²⁾; logischerweise müsste man hieraus auf eine grössere Geburtenhäufigkeit schliessen; da nun aber gerade das Gegenteil stattfindet, so bleibt nur der Schluss auf ein rapides Anwachsen der nicht registrierpflichtigen Geburten aller Grade, verbunden mit einem Überhandnehmen der konzeptionsverhindernden Mittel.

4. Zu interessanten Vergleichen führt folgende Aufgabe: Von einem Kollektivgegenstand gegebenen Umfangs liegen zwei veränderliche Merkmale vor; gegeben sind ferner die Häufigkeiten für das Auftreten jedes einzelnen dieser Merkmale sowie für ihr gleichzeitiges Auftreten. Welche Bedingung muss erfüllt sein, damit Unabhängigkeit zwischen dem Auftreten der beiden Merkmale angenommen werden kann? Wir wählen wiederum ein Beispiel aus der schweizerischen Statistik. In der Stadt Zürich verteilten sich die *Eheschliessenden* des Jahres 1920 auf die verschiedenen *Altersklassen* wie folgt³⁾:

¹⁾ Zeitschrift des königl. sächsischen statistischen Landesamtes 1916/17, S. 28.

²⁾ G. v. Bunge: Die zunehmende Unfähigkeit der Frauen, ihre Kinder zu stillen, die Ursachen dieser Unfähigkeit, die Mittel zur Verhütung, 7. A., München 1914.

³⁾ Statistisches Jahrbuch der Stadt Zürich, 1920 und 1921, S. 7.

Alter des Mannes 3 Jahre	Alter der Frau, Jahre											Total Männer
	bis 20	20-25	25-30	30-35	35-40	40-45	45-50	50-55	55-60	60-65	über 65	
bis 20	2	4	—	—	—	—	—	—	—	—	—	6
20-25	31	298	170	34	5	2	—	—	—	—	—	540
25-30	19	364	423	85	24	5	1	1	—	1	—	923
30-35	5	156	207	97	41	10	3	—	—	—	—	519
35-40	4	33	87	72	36	12	6	2	—	—	—	252
40-45	1	13	34	41	30	17	11	1	—	—	—	148
45-50	1	6	8	27	21	20	8	1	—	1	—	93
50-55	—	2	8	6	9	9	8	11	—	1	1	55
55-60	—	1	—	5	3	9	4	11	3	3	1	40
60-65	—	1	—	3	2	5	6	4	1	4	—	26
über 65	—	1	—	—	1	3	3	—	2	2	—	12
Total Frauen. .	63	879	937	370	172	92	50	31	6	12	2	2614

Zwei Grössen heissen voneinander unabhängig, wenn eine Veränderung der einen keine Veränderung der andern zur Folge hat. Nun hat man es hier nicht mit einer funktionalen, sondern mit einer korrelativen Abhängigkeit zu tun. Diese letztere kann auf die erste durch Einführung von Mittelwerten zurückgeführt werden; einen andern Weg zur Erfassung korrelativer Abhängigkeiten gibt es nicht. Wir heissen die Alter der Heiratenden voneinander unabhängig, wenn zu einem veränderlichen Alter des Mannes ein unveränderliches mittleres Alter der Frau und umgekehrt zu einem veränderlichen Alter der Frau ein unveränderliches mittleres Alter des Mannes gehört. Dies ist dann der Fall, wenn die prozentuale Besetzung aller Zeilen dieselbe ist und wenn das gleiche für die prozentuale Besetzung der Kolonnen gilt. Die verlangte Unabhängigkeit wird darum durch die folgende Tabelle erfüllt:

Alter des Mannes 4 Jahre	Alter der Frau, Jahre											Total Männer
	bis 20	20-25	25-30	30-35	35-40	40-45	45-50	50-55	55-60	60-65	über 65	
bis 20	—	2	2	1	1	—	—	—	—	—	—	6
20-25	13	182	194	77	35	19	10	6	1	3	—	540
25-30	22	310	331	130	61	33	18	11	2	4	1	923
30-35	13	176	187	74	34	18	10	6	1	—	—	519
35-40	6	84	90	35	16	9	5	3	1	2	1	252
40-45	4	50	53	21	10	5	3	2	—	—	—	148
45-50	2	31	34	13	6	4	2	1	—	—	—	93
50-55	1	18	19	8	3	2	1	1	1	1	—	55
55-60	1	13	14	5	3	1	1	1	—	1	—	40
60-65	1	9	9	4	2	1	—	—	—	—	—	26
über 65	—	4	4	2	1	—	—	—	—	1	—	12
Total Frauen. .	63	879	937	370	172	92	50	31	6	12	2	2614

Man wird bemerken, dass sich die Besetzung irgendeines Feldes, z. B. der Kombination 30-35, 45-50, durch Multiplikation der Wahrscheinlichkeiten

für das Auftreten der kombinierten Altersklassen mit dem Umfang des Kollektivs (2614) ergibt. In unserm Beispiel ist die Wahrscheinlichkeit für das Auftreten der Altersklasse 30—35 beim Manne: $\frac{519}{2614}$ und die Wahrscheinlichkeit für

das Auftreten der Altersklasse 45—50 bei der Frau: $\frac{50}{2614}$; in der Tat ist

$$\frac{519}{2614} \cdot \frac{50}{2614} \cdot 2614 = \frac{25.950}{2614} = 10, \text{ wie es sein soll.}$$

Man überzeugt sich, dass diese Regel sich deckt mit derjenigen für das gleichzeitige Auftreten zweier voneinander unabhängiger Ereignisse, falls deren Wahrscheinlichkeiten gegeben sind. Wir können die Häufigkeit zweier kombinierter Merkmale, für den Fall der Unabhängigkeit, als normale Häufigkeit bezeichnen, weil sie dem normal zufälligen Geschehen entspricht. Demnach enthält die letzte Tabelle normale Häufigkeiten für kombinierte Heiratsalter. Drückt man die Differenz zwischen der tatsächlichen und normalen Häufigkeit zweier kombinierten Merkmale in Prozenten der normalen Häufigkeit aus, so misst man den *Abhängigkeitsgrad* beider Merkmale. Je nachdem diese Prozentzahlen positiv oder negativ sind, spricht man von *positiver* oder *negativer Abhängigkeit*; der höchste Grad negativer Abhängigkeit wird demnach durch -100 dargestellt; die positive Abhängigkeit hat überhaupt keine obere Grenze, ihre untere Grenze ist natürlich die Null, das Symbol der Unabhängigkeit. Nach diesen Grundsätzen ist die folgende Tabelle entstanden. Sie misst die Abhängigkeit der kombinierten Heiratsalter für die Eheschliessenden der Stadt Zürich im Jahre 1920:

Alter des Mannes 5 Jahre	Alter der Frau, Jahre										
	bis 20	20—25	25—30	30—35	35—40	40—45	45—50	50—55	55—60	60—65	über 65
bis 20	+ 100	— 100	— 100	— 100
20—25 . . .	+ 138	+ 64	— 12	— 56	— 86	— 89	— 100	— 100	— 100	— 100	.
25—30 . . .	— 14	+ 17	+ 28	— 35	— 61	— 85	— 94	— 91	— 100	— 75	— 100
30—35 . . .	— 62	— 11	+ 11	+ 31	— 21	— 44	— 70	— 100	— 100	.	.
35—40 . . .	— 33	— 61	— 3	+ 106	+ 125	+ 33	+ 20	— 33	— 100	— 100	— 100
40—45 . . .	— 75	— 74	— 36	+ 95	+ 200	+ 240	+ 267	— 50	.	.	.
45—50 . . .	— 50	— 81	— 76	+ 108	+ 250	+ 400	+ 300
50—55 . . .	— 100	— 89	— 58	— 25	+ 200	+ 350	+ 700	+ 1000	— 100	.	.
55—60 . . .	— 100	— 92	— 100	.	.	+ 800	+ 300	+ 1000	.	+ 200	.
60—65 . . .	— 100	— 89	— 100	— 25	.	+ 400
über 65	— 100	— 100	+ 100	.

Trotz der kleinen Beobachtungszahl kann man deutlich erkennen, dass eine geschlossene Linie die positiven Abhängigkeiten von den negativen trennt; man kann eine ganze Schar ähnlicher Linien erhalten, indem die Punkte gleicher Abhängigkeitsgrade miteinander verbunden werden; Voraussetzung hierzu ist allerdings eine grosse Beobachtungszahl und eine graphische Darstellung, so dass in einem rechtwinkligen Koordinatensystem das Alter der Frau als Abszisse und dasjenige des Mannes als Ordinate aufgetragen wird. Diese in Prozenten

ausgedrückten Abhängigkeitsgrade geben an, um wie viele Prozente die beobachteten Häufigkeiten grösser oder kleiner sind als die normalen. Addiert man zu jeder Zahl 100 %, so erhält man die beobachteten Häufigkeiten, ausgedrückt in Prozenten der normalen. Bei dieser Darstellung bedeutet 100 % Unabhängigkeit zwischen den beiden Merkmalen.

Wir wollen in gleicher Weise die Abhängigkeitsgrade zwischen der *Tuberkulosensterblichkeit* und den verschiedenen Altersgruppen ermitteln. Wir beschränken uns auf die Sterbefälle *männlichen* Geschlechts in der Schweiz im Jahre 1924. Die folgende Tabelle gibt den gewünschten Aufschluss ¹⁾:

6	0	1-11	1	2-4	5-14	15-19	20-29	30-39	40-49	50-59	60-69	70-79	80 und mehr	Total
1. Altersgruppen	Mo-nat	Mo-nate	Jahr	Jahre										
2. Sterbefälle im ganzen . . .	1431	1180	341	406	569	454	1253	1316	2305	3575	4822	4568	2249	24469
3. Sterbefälle durch Tuberkulose . . .	—	46	35	59	111	152	566	453	477	445	296	109	18	2767
4. Tuberkulosensterbefälle bei normaler Verteilung . . .	162	133	39	46	64	51	142	149	261	404	545	517	254	2767
5. Differenzen zwischen 3.u.4.	-162	-87	+4	+13	+47	+101	+424	+304	+216	+41	-249	-408	-236	—
6. Abhängigkeit vom Alter = $\frac{3.-4.}{4.} \cdot 100$	-100	-65	+10	+28	+73	+198	+299	+204	+83	+10	-46	-79	-93	—

Linie 6 beweist, dass zwischen Alter und Tuberkulosensterblichkeit für das männliche Geschlecht das Maximum der positiven Abhängigkeit in die Altersgruppe 20—29 fällt, während die extremen Alter auch eine extreme negative Abhängigkeit zeigen. Die Altersgruppe 20—29 hat eine viermal grössere Sterblichkeit an Tuberkulose, als nach der normalen Verteilung der Tuberkulosensterbefälle (welche also prozentual mit der Verteilung aller männlichen Todesfälle übereinstimmt) zu erwarten wäre.

In ähnlicher Weise wollen wir die Abhängigkeit der *Krebssterblichkeit* vom Alter feststellen. Wir benützen wiederum die Todesfälle *männlichen* Geschlechts in der Schweiz für das Jahr 1924. Die umstehenden Zahlen (Tab. 7) enthalten die Resultate.

Das Maximum der positiven Abhängigkeit beträgt beim Krebs bloss 84 %, und zwar in der Altersgruppe 60—69; die extreme negative Abhängigkeit hält an vom Alter 0 bis zur Altersgruppe 20—29 (inklusive). Indem wir zusammenfassen, erhalten wir die folgenden Gegenüberstellungen (siehe Tab. 8 und 9).

Der Vollständigkeit halber teilen wir die entsprechenden Zahlen auch für das *weibliche* Geschlecht mit (siehe Tab. 10 und 11).

¹⁾ Schweizerische statistische Mitteilungen, 1927, 1. Heft.

7 1. Altersgruppen	0 Mo- nat	1—11 Mo- nate	1 Jahr	2—4 Jahre	5—14 Jahre	15—19 Jahre	20—29 Jahre	30—39 Jahre	40—49 Jahre	50—59 Jahre	60—69 Jahre	70—79 Jahre	80 und mehr Jahre	Total
2. Sterbefälle im ganzen . . .	1431	1180	341	406	569	454	1253	1316	2305	3575	4822	4568	2249	24469
3. Sterbefälle durch Krebs .	—	1	—	—	—	1	11	47	268	683	1017	655	128	2811
4. Krebssterbe- fälle bei nor- maler Ver- teilung . . .	164	136	39	47	65	52	144	151	265	411	554	525	258	2811
5. Differenzen zwischen 3. u. 4.	-164	-135	-39	-47	-65	-51	-133	-104	+3	+272	+463	+130	-130	—
6. Abhängigkeit vom Alter = $\frac{3.-4.}{4.} \cdot 100$	-100	-99	-100	-100	-100	-98	-92	-69	+1	+66	+84	+25	-50	—

Tuberkulosensterblichkeit in der Schweiz 1924, männliches Geschlecht

8 Altersgruppen	0 Mo- nat	1—11 Mo- nate	1 Jahr	2—4 Jahre	5—14 Jahre	15—19 Jahre	20—29 Jahre	30—39 Jahre	40—49 Jahre	50—59 Jahre	60—69 Jahre	70—79 Jahre	80 und mehr Jahre
Sterben bei normaler Verteilung der Todesfälle . . .	100	100	100	100	100	100	100	100	100	100	100	100	100
so sterben nach Be- obachtung . . .	0	35	90	128	173	298	399	304	183	110	54	21	7

Krebssterblichkeit in der Schweiz 1924, männliches Geschlecht

9 Altersgruppen	0 Mo- nat	1—11 Mo- nate	1 Jahr	2—4 Jahre	5—14 Jahre	15—19 Jahre	20—29 Jahre	30—39 Jahre	40—49 Jahre	50—59 Jahre	60—69 Jahre	70—79 Jahre	80 und mehr Jahre
Sterben bei normaler Verteilung der Todesfälle . . .	100	100	100	100	100	100	100	100	100	100	100	100	100
so sterben nach Be- obachtung . . .	0	1	0	0	0	2	8	31	101	166	184	125	50

10 1. Altersgruppen	0 Mo- nat	1-11 Mo- nate	1 Jahr	2-4 Jahre	5-14 Jahre	15-19 Jahre	20-29 Jahre	30-39 Jahre	40-49 Jahre	50-59 Jahre	60-69 Jahre	70-79 Jahre	80 und mehr Jahre	Total
2. Sterbefälle im ganzen . . .	1048	900	296	338	471	557	1442	1355	1820	2888	4595	5634	3175	24519
3. Sterbefälle durch Tuberkulose	—	37	39	68	138	342	876	481	367	348	310	190	31	3227
4. Tuberkulosensterbefälle bei normaler Verteilung . . .	138	118	39	45	62	73	190	178	240	380	605	741	418	3227
5. Differenzen zwischen 3. u. 4.	-138	-81	0	+23	+76	+269	+686	+303	+127	-32	-295	-551	-387	—
6. Abhängigkeit vom Alter = $\frac{3.-4.}{4.} \cdot 100$. . .	-100	-69	0	+51	+123	+368	+361	+170	+53	-8	-49	-74	-93	—

11 1. Altersgruppen	0 Mo- nat	1-11 Mo- nate	1 Jahr	2-4 Jahre	5-11 Jahre	15-19 Jahre	20-29 Jahre	30-39 Jahre	40-49 Jahre	50-59 Jahre	60-69 Jahre	70-79 Jahre	80 und mehr Jahre	Total
2. Sterbefälle im ganzen . . .	1048	900	296	338	471	557	1442	1355	1820	2888	4595	5634	3175	24519
3. Sterbefälle durch Krebs .	1	—	—	1	—	1	5	90	280	609	787	632	204	2610
4. Krebssterbefälle bei normaler Verteilung . . .	112	96	32	36	50	59	153	144	194	307	489	600	338	2610
5. Differenzen zwischen 3. u. 4.	-111	-96	-32	-35	-50	-58	-148	-54	+86	+302	+298	+32	-134	—
6. Abhängigkeit vom Alter = $\frac{3.-4.}{4.} \cdot 100$. . .	-99	-100	-100	-97	-100	-98	-97	-37	+44	+98	+61	+5	-40	—

Durch Zusammenfassung erhalten wir folgendes Bild:

Tuberkulosensterblichkeit in der Schweiz 1924, weibliches Geschlecht

12 Altersgruppen	0 Mo- nat	1-11 Mo- nate	1 Jahr	2-4 Jahre	5-14 Jahre	15-19 Jahre	20-29 Jahre	30-39 Jahre	40-49 Jahre	50-59 Jahre	60-69 Jahre	70-79 Jahre	80 und mehr Jahre
Sterben bei normaler Verteilung der Todesfälle	100	100	100	100	100	100	100	100	100	100	100	100	100
so sterben nach Beobachtung	0	31	100	151	223	468	461	270	153	92	51	26	7

Krebssterblichkeit in der Schweiz 1924, weibliches Geschlecht

13 Altersgruppen	0 Mo- nat	1—11 Mo- nate	1 Jahr	2—4 Jahre	5—14 Jahre	15—19 Jahre	20—29 Jahre	30—39 Jahre	40—49 Jahre	50—59 Jahre	60—69 Jahre	70—79 Jahre	80 und mehr Jahre
Sterben bei nor- maler Verteilung der Todesfälle .	100	100	100	100	100	100	100	100	100	100	100	100	100
so sterben nach Beobachtung . .	1	0	0	3	0	2	3	63	144	198	161	105	60

Der Vergleich für beide Geschlechter zeigt, dass bis und mit der Altersgruppe 20—29 das weibliche Geschlecht eine relativ grössere Tuberkulosensterblichkeit besitzt als das männliche; für die folgenden Altersgruppen besteht das umgekehrte Verhältnis. Bei der Krebssterblichkeit beobachten wir dieselbe Erscheinung, nur kehrt sich das Verhältnis erst mit der Altersgruppe 60—69 um.

Die Vergleichszahlen der beiden letzten Zusammenstellungen sowie die entsprechenden für das männliche Geschlecht kann man direkt folgendermassen berechnen: Man ermittelt für das betreffende Geschlecht die prozentuale Verteilung aller Todesfälle nach dem *Alter*, tut dasselbe für die Tuberkulosen- bzw. Krebssterbefälle; sodann dividiert man die Prozentzahlen der zweiten Reihe durch die entsprechenden (d. h. zur gleichen Altersgruppe gehörigen) der ersten und multipliziert jeden der erhaltenen Quotienten mit 100.

Zu dem Begriff der *positiven und negativen Abhängigkeit* gelangten wir durch Zugrundelegung von Mittelwerten, er erhält seine Rechtfertigung durch diese. Yule hat zur Messung solcher Abhängigkeiten den sogenannten *Abhängigkeitskoeffizienten* eingeführt; derselbe bewegt sich zwischen den Grenzen -1 und $+1$, hat aber den Nachteil, dass aus ihm das Verhältnis zwischen normaler und beobachteter Häufigkeit nicht ohne weiteres abgelesen werden kann; zudem kann man bei denselben Grenzen auf verschiedene Arten solche Koeffizienten bilden, am einfachsten, indem die Differenz zwischen normaler und beobachteter Häufigkeit durch deren Summe dividiert wird. Aus diesen Gründen haben wir von dem Yuleschen Abhängigkeitskoeffizienten keinen Gebrauch gemacht ¹⁾.

5. Anlass zu interessanten Vergleichen geben die *Alterskombinationen der zusammenlebenden Ehegatten*. Da unseres Wissens für die schweizerischen Verhältnisse keine bezüglichen Untersuchungen vorliegen, so sei es gestattet, hierauf etwas näher einzutreten. Wir wählen zunächst die bezüglichen Zahlen der Volkszählung von 1910, verfahren aber repräsentativ, indem bloss Alter im Abstand von fünf zu fünf Jahren berücksichtigt werden. Die Ausgangszahlen sind die folgenden ²⁾:

¹⁾ Zu den angestellten Vergleichen hat mich veranlasst das Werk von Dr. Emanuel Czuber: Die statistischen Forschungsmethoden, Wien 1921, S. 10 ff. Man findet hier auch viele Literaturangaben.

²⁾ Die Ergebnisse der eidgenössischen Volkszählung vom 1. Dezember 1910, II. Bd., S. 396.

Alterskombination der zusammenlebenden einheimischen und ausländischen Ehegatten im Abstand von fünf zu fünf Jahren, 1910

14 Alter der Frau	Alter des Mannes														Reihentotal	Mittleres Alter des Mannes	
	22	27	32	37	41	47	52	57	62	67	72	77	82	87			92
17	14	6	2	—	—	—	—	—	—	—	—	—	—	—	—	22	24,3
22	311	676	262	67	26	10	—	1	1	—	—	—	—	—	—	1354	27,8
27	79	1502	1244	409	139	41	13	4	2	3	1	—	—	—	—	3437	30,9
32	12	320	1791	1315	481	165	50	15	7	5	2	—	—	—	1	4164	35,4
37	3	67	382	1627	998	454	142	52	13	9	2	1	—	—	—	3750	39,6
42	2	18	97	388	1237	944	344	117	44	20	4	2	—	—	—	3217	44,6
47	—	6	34	92	334	1153	744	329	105	43	10	10	1	—	—	2861	49,4
52	—	1	7	22	68	282	751	558	248	88	36	9	1	—	—	2071	54,4
57	—	—	2	5	23	62	171	605	389	209	60	22	—	—	2	1550	59,2
62	—	—	1	1	5	20	48	138	372	295	122	31	7	1	—	1041	63,6
67	—	—	—	—	2	4	25	38	98	267	177	45	15	7	—	678	67,5
72	—	—	—	—	1	1	3	5	14	59	133	87	23	4	—	330	72,3
77	—	—	—	—	—	—	—	2	4	11	18	56	23	2	1	117	75,7
82	—	—	—	—	—	—	—	—	1	2	4	5	10	4	—	26	78,3
87	—	—	—	—	—	—	—	—	—	—	1	1	2	1	1	6	82,0
Total	421	2596	3822	3926	3314	3136	2291	1864	1298	1011	570	269	82	22	2	24624	43,8
Mittleres Alter der Frau																	
	23,3	26,7	30,6	34,6	39,1	43,7	48,1	52,6	56,7	60,7	64,7	68,1	72,7	69,7	82,0	40,9	

Die letzte Kolonne enthält das mittlere Alter des Mannes zu einem gegebenen Alter der Frau und die letzte Zeile das mittlere Alter der Frau zu einem gegebenen Alter des Mannes. Alle Männer zusammen besitzen das mittlere Alter 43,8; alle Frauen zusammen sind im Mittel 40,9 Jahre alt. Die ausgewählten Alterskombinationen weisen auf ausgesprochene Korrelation zwischen dem Alter des Mannes und demjenigen der Frau hin. Die Frage liegt nun nahe, ob dieses *Repräsentativverfahren* annähernd dieselbe Korrelation liefert wie bei Berücksichtigung sämtlicher Alter. Eine solche Korrelation wird bestimmt durch die verzeichneten mittleren Alter des Mannes und diejenigen der Frau sowie durch die zugehörigen Gewichte, gegeben im Linientotal bzw. Kolonnentotal. Nun soll uns der nachfolgende Auszug aus der letzten Tabelle leiten:

15														
Alter des Mannes														
22	27	32	37	42	47	52	57	62	67	72	77	82	87	92
Mittleres Alter der Frau														
23,3	26,7	30,6	34,9	39,1	43,7	48,1	52,6	56,7	60,7	64,7	68,1	72,7	69,7	82,0
Mittlere Altersdifferenz														
-1,3	+0,3	+1,4	+2,1	+2,9	+3,3	+3,9	+4,4	+5,3	+6,3	+7,3	+8,9	+9,3	+17,3	+10,0

Diesen Zahlen stellen wir die folgenden aus dem 2. Band zur Volkszählung von 1910 gegenüber ¹⁾; die eingeklammerten Zahlen haben wir der bessern Vergleichsfähigkeit halber zugefügt:

16															Alter des Mannes														
20-24	25-29	30-34	35-39	40-44	45-49	50-54	55-59	60-64	65-69	70-74	75-79	80-84	85-89	90-94															
Annäherndes Mittel																													
(22)	(27)	(32)	(37)	(42)	(47)	(52)	(57)	(62)	(67)	(72)	(77)	(82)	(87)	(92)															
Mittleres Alter der Frau																													
(23,1)	(26,6)	(30,5)	(34,7)	(39,2)	(43,6)	(48,0)	(52,4)	(56,8)	(60,8)	(64,5)	(68)	(70,8)	(73,4)	.															
Mittlere Altersdifferenzen																													
-1,09	+0,40	+1,49	+2,28	+2,84	+3,41	+4,01	+4,62	+5,21	+6,22	+7,53	+8,99	+11,24	+13,65	.															

Durch Vergleich findet man, dass das Repräsentativverfahren fast dasselbe mittlere Frauenalter ergibt, wie die umfassenderen Berechnungen, abgesehen von den Grenzaltern, die aber zufolge ihrer schwachen Besetzung wenig ins Gewicht fallen. Das gleiche gilt natürlich auch für den Unterschied zwischen dem Alter des Mannes und dem mittleren Alter der Frau; aber auch die zugehörigen Gewichtszahlen zeigen denselben Grad von Übereinstimmung, wie die folgenden Relativzahlen beweisen:

17																	Alter des Mannes																
22	27	32	37	42	47	52	57	62	67	72	77	82	87	92	Total																		
Gewichte = Kolonnensumme in % der Totalsumme																																	
1,7	10,5	15,5	16,0	13,5	12,7	9,3	7,6	5,3	4,1	2,3	1,1	0,3	0,1	0	100																		

18																		Alter des Mannes																	
20-24	25-29	30-34	35-39	40-44	45-49	50-54	55-59	60-64	65-69	70-74	75-79	80-84	85-89	90-94	Total																				
Gewichte = Kolonnensummen in % der Totalsumme																																			
2,0	10,0	15,5	15,6	13,7	12,3	9,9	7,4	5,6	4,2	2,4	1,0	0,3	0,1	0	100																				

¹⁾ S. 42*.

Für das mittlere Alter des Mannes zu einem gegebenen Alter der Frau und dem zugehörigen Gewicht sind die Verhältnisse ganz analog. Wir können also die gestellte Frage nach der Güte des gewählten Repräsentativverfahrens bejahen und damit feststellen, dass die Differenz zwischen dem Alter des Mannes und dem zugehörigen mittleren Alter der Frau mit wachsendem Alter des Mannes ebenfalls wächst. Das umgekehrte findet statt für die Differenz zwischen dem mittleren Alter des Mannes und dem zugehörigen Alter der Frau, wie die folgende Zusammenstellung zeigt:

19														
Alter der Frau														
17	22	27	32	37	42	47	52	57	62	67	72	77	82	87
Mittleres Alter des Mannes														
24,3	27,8	30,9	35,4	39,9	44,6	49,4	54,4	59,2	63,6	67,5	72,3	75,7	78,3	82,0
Mittlere Altersdifferenz														
+7,3	+5,8	+3,9	+3,4	+2,9	+2,6	+2,4	+2,4	+2,2	+1,6	+0,5	+0,3	-1,3	-3,7	-5,0

Wählt man, zwecks *graphischer Darstellung*, ein rechtwinkliges Koordinatensystem und trägt in diesem das Alter des Mannes als Abszisse und das zugehörige der Frau als Ordinate auf, so entspricht jedem Ordinateneckpunkt eine bestimmte Alterskombination; trägt man zu jedem Alter des Mannes die zugehörigen mittleren Alter der Frau als Ordinate auf, so wird man bemerken, dass die erhaltenen Ordinateneckpunkte (man heisst sie die repräsentierenden Punkte) sich gut einer Geraden anpassen, abgesehen von den äussersten Altern, die aber der schwachen Besetzung wegen wenig ins Gewicht fallen. Es liegt nun nahe, diese repräsentierenden Punkte durch eine Gerade zu ersetzen, welche sich denselben am besten anschmiegt; dies kann graphisch, mit Hilfe eines Fadens geschehen oder aber rechnerisch mit Hilfe der Methode der kleinsten Quadrate ¹⁾; wir haben das letzte Verfahren, als das sicherste, gewählt. Man lässt die Gerade durch den

20															
Alter des Mannes															
22	27	32	37	42	47	52	57	62	67	72	77	82	87	92	
Mittleres Alter der Frau nach der 1. Korrelationsachse															
22,3	26,6	30,8	35,1	39,4	43,6	47,9	52,2	56,4	60,7	65,0	69,2	73,5	77,8	82,1	
Mittlere Altersdifferenzen															
-0,3	+0,4	+1,2	+1,9	+2,6	+3,4	+4,1	+4,8	+5,6	+6,3	+7,0	+7,8	+8,5	+9,2	+9,9	

¹⁾ Czuber, op. cit., S. 113 ff.

jenigen Punkt hindurch gehen, welcher das mittlere Alter aller Männer (43,8) und dasjenige aller Frauen (40,9) repräsentiert. Man kann diesen Punkt als Schwerpunkt oder Mittelpunkt der Korrelationstabelle bezeichnen. Nach der Bestimmung dieser Geraden (1. Korrelationsachse) konstatiert man, dass beim Fortschreiten auf ihr das mittlere Alter der Frau um 8,53 Jahre steigt oder sinkt, je nachdem das Alter des Mannes um 10 Jahre zu- oder abnimmt; damit im Einklang stehen die Zahlen der Tab. 20 (oben):

Diese Zahlen stimmen (mit Ausschluss der äussersten Alter) gut mit den entsprechenden in Tab. 15 und Tab. 16 überein.

Bestimmt man in analoger Weise diejenige Gerade durch den Schwerpunkt, welche sich den repräsentierenden Punkten der Zeilenmittel am besten anpasst, so findet man, dass beim Fortschreiten auf dieser zweiten Korrelationsachse das mittlere Alter des Mannes um 9,19 Jahre steigt oder sinkt, je nachdem das Alter der Frau um 10 Jahre zu- oder abnimmt, in Übereinstimmung mit den folgenden Zahlen:

21		Alter der Frau													
17	22	27	32	37	42	47	52	57	62	67	72	77	82	87	
Mittleres Alter des Mannes nach der 2. Korrelationsachse															
21,8	26,4	31,0	35,6	40,2	44,8	49,4	54,0	58,6	63,2	67,8	72,4	77,0	81,6	86,2	
Mittlere Altersdifferenzen															
+4,8	+4,4	+4,0	+3,6	+3,2	+2,8	+2,4	+2,0	+1,6	+1,2	+0,8	+0,4	0	-0,4	-0,8	

Die Übereinstimmung mit der entsprechenden Tabelle 19 ist auch hier eine gute, sofern man von den extremen Altern absieht.

Nun untersuchen wir die Alterskombinationen der zusammenlebenden Ehepaare auf Grund der Volkszählung von 1920, indem wir dasselbe Repräsentativverfahren anwenden, wie für 1910. Wir gehen von den folgenden Zahlen aus 1).

Alterskombination der zusammenlebenden einheimischen und ausländischen Ehegatten, im Abstand von fünf zu fünf Jahren, 1920

Alter der Frau	22															Reihentotal	Mittleres Alter des Mannes
	Alter des Mannes																
	22	27	32	37	42	47	52	57	62	67	72	77	82	87	92		
17	—	5	1	—	—	—	—	—	—	—	—	—	—	—	—	6	27,8
22	330	589	234	67	23	7	3	3	—	—	—	—	—	—	—	1256	27,7
Übertrag	330	594	235	67	23	7	3	3	—	—	—	—	—	—	—	1262	

1) Schweizerische statistische Mitteilungen, 1925. 6. Heft, S. 54.

Alter der Frau	Alter des Mannes														Reihentotal	Mittleres Alter des Mannes	
	22	27	32	37	42	47	52	57	62	67	72	77	82	87			92
Übertrag	330	594	235	67	23	7	3	3	—	—	—	—	—	—	—	1262	
27	71	1441	1207	420	155	43	18	7	4	—	—	—	—	—	—	3366	31,1
32	13	247	1604	1313	471	135	48	21	5	4	—	—	—	—	—	3861	35,6
37	—	56	356	1707	1264	453	150	56	15	6	1	2	—	—	—	4066	40,1
42	1	11	83	385	1624	1149	437	143	48	18	3	1	2	—	—	3905	44,8
47	1	10	27	80	333	1357	847	392	104	35	7	4	—	—	—	3197	49,4
52	—	2	13	21	92	342	1029	741	272	79	31	10	—	—	—	2632	54,0
57	—	—	—	7	29	68	276	841	526	218	57	17	2	1	—	2042	58,7
62	—	—	2	2	5	20	44	198	472	309	136	32	8	2	—	1230	63,3
67	—	1	—	—	6	4	24	39	109	327	167	67	8	4	—	756	67,2
72	—	—	—	—	—	—	4	9	22	55	129	82	18	2	1	322	71,8
77	—	—	—	—	—	—	2	5	12	11	19	51	22	2	1	125	73,9
82	—	—	—	—	—	—	—	—	—	1	1	6	8	4	—	21	79,1
87	—	—	—	—	—	—	—	—	—	—	2	—	—	2	—	4	79,5
Total	416	2362	3527	4002	4002	3578	2882	2456	1589	1063	553	272	68	17	2	26789	44,7
Mittleres Alter der Frau																	
23,3 26,7 30,6 34,9 39,4 44,0 48,5 52,9 57,2 61,4 65,1 68,3 71,7 73,5 74,5 41,9																	

Aus dieser Tabelle, welche vollständig gleich wie Tabelle 14 aufgebaut ist (also keiner weitem Erläuterung bedarf), geht die folgende hervor:

23														
Alter des Mannes														
22	27	32	37	42	47	52	57	62	67	72	77	82	87	92
Mittleres Alter der Frau														
23,3	26,7	30,6	34,9	39,4	44,0	48,5	52,9	57,2	61,4	65,1	68,3	71,7	73,5	74,5
Mittlere Altersdifferenzen														
-1,3	+0,3	+1,4	+2,1	+2,6	+3,0	+3,5	+4,1	+4,8	+5,6	+6,9	+8,7	10,3	13,5	17,5

Vergleicht man diese Zahlen mit Tabelle 15, so findet man, dass seit 1910 die Altersdifferenzen abgenommen haben. Als Vergleichsstück zu Tabelle 19 erhalten wir die Zahlen der Tabelle 24 (umstehend).

Der Vergleich mit Tabelle 19 zeigt, dass seit 1910 die Differenz zwischen dem Alter der Frau und dem mittleren Alter des Mannes bis zum Alter 47 zugenommen hat, um von da an wieder abzunehmen. Dieselben Feststellungen kann man bestätigen, wenn für die mittleren Alter (Zeilenmittel und Kolonnenmittel) wiederum Mittelwerte verwendet werden; hierzu werden wir durch den Umstand

24														
Alter der Frau														
17	22	27	32	37	42	47	52	57	62	67	72	77	82	87
Mittleres Alter des Mannes														
27,8	27,7	31,1	35,8	40,1	44,8	49,4	54,0	58,7	63,3	67,2	71,8	73,9	79,1	79,5
Mittlere Altersdifferenzen														
+10,8	+5,7	+4,1	+3,8	+3,1	+2,8	+2,4	+2,0	+1,7	+1,3	+0,2	-0,2	-3,1	-2,9	-7,8

ermutigt, dass auch hier, bei graphischer Darstellung des mittleren Frauenalters zu einem gegebenen Alter des Mannes und umgekehrt, die repräsentierenden Punkte in der Nähe einer geraden Linie liegen; die äussersten Alter weichen hiervon stärker ab, was aber bei ihrer schwachen Besetzung wenig ins Gewicht fällt. Bestimmt man nun, nach der Methode der kleinsten Quadrate, diejenigen Geraden durch den Schwerpunkt (44,7/41,9), welche die beste Anpassung an die repräsentierenden Punkte zeigen, so erhält man die folgenden Resultate:

Beim Fortschreiten auf der einen dieser Geraden (1. Korrelationsachse) steigt oder sinkt das mittlere Alter der Frau um 8,69 Jahre, je nachdem das zugehörige Alter des Mannes um 10 Jahre zu- oder abnimmt;

Beim Fortschreiten auf der andern dieser beiden Geraden (2. Korrelationsachse) steigt oder sinkt das mittlere Alter des Mannes um 9,04 Jahre, je nachdem das zugehörige Alter der Frau um 10 Jahre zu- oder abnimmt. Wir lassen die eingehenden zahlenmässigen Darstellungen folgen:

25														
Alter des Mannes														
22	27	32	37	42	47	52	57	62	67	72	77	82	87	92
Mittleres Alter der Frau nach der 1. Korrelationsachse														
22,2	26,5	30,9	35,2	39,8	43,9	48,2	52,6	56,9	61,3	65,8	69,9	74,3	78,8	83,0
Mittlere Altersdifferenzen														
-0,2	+0,5	+1,1	+1,8	+2,4	+3,1	+3,8	+4,4	+5,1	+5,7	+6,4	+7,1	+7,7	+8,4	+9,0

Der Vergleich von Tabelle 25 mit Tabelle 20 einerseits und derjenige von Tabelle 26 (umstehend) mit Tabelle 21 andererseits, führt zu den bereits bekannten Resultaten.

Der soeben geführte Vergleich zwischen zwei Systemen korrelativer Abhängigkeiten stützt sich auf *Mittelwertbildungen* und teilt darum alle Vor- und Nachteile mit diesen; ohne Mittelwerte und Vergleichen ist keine wissenschaftliche Statistik denkbar; denn wissenschaftliches Arbeiten setzt Vergleichen voraus; sodann ist der Mittelwertbegriff nicht von demjenigen der

26														Alter der Frau		
17	22	27	32	37	42	47	52	57	62	67	72	77	82	87		
Mittleres Alter des Mannes nach der 2. Korrelationsachse																
22,2	26,7	31,2	35,8	40,3	44,8	49,3	53,8	58,3	62,9	67,4	71,9	76,4	81,0	85,5		
Mittlere Altersdifferenzen																
+5,2	+4,7	+4,2	+3,8	+3,3	+2,8	+2,3	+1,8	+1,3	+0,9	+0,4	-0,1	-0,6	-1,0	-1,5		

Statistik zu trennen, weil man typische Eigentümlichkeiten der sozialen Erscheinungen ohne Mittelwerte nicht erfassen kann. Ein Typus ist immer an Mittelwerte gebunden. Wir haben die zu den mittleren Altern gehörenden repräsentierenden Punkte durch gerade Linien ersetzt; die Berechtigung hierzu kann ernstlich bezweifelt werden, wir haben darum diese Annahme, welche die Einführung von neuen Mittelwerten bedeutet, einer besondern Prüfung unterzogen, indem die Richtungen der Korrelationsachsen auf zwei verschiedene Arten bestimmt wurden; die Resultate deckten sich vollständig, was durchaus zugunsten der umstrittenen Annahme spricht. Zu genau demselben Ergebnis gelangt man durch Verwendung des *Korrelationsverhältnisses* von Pearson ¹⁾. Es wird nicht ganz überflüssig sein, die Elemente zusammenzustellen, welche den Vergleich der beiden Korrelationstabellen ermöglichen ²⁾:

27	Schwerpunktskoordinaten	Winkel der 1. Korrelationsachse mit der Abszissenachse	Tangente dieses Winkels	Winkel der 2. Korrelationsachse mit der Ordinatenachse	Tangente dieses Winkels	Winkel der beiden Korrelationsachsen
Korrelationstabelle von 1910 . . .	43,8 40,9	40° 28'	0,853	42° 35'	0,919	6° 57'
Korrelationstabelle von 1920 . . .	44,7 41,9	40° 59'	0,869	42° 7'	0,904	6° 54'

¹⁾ Czuber, op. cit., S. 143.

²⁾ Ich bin über die Herleitung dieser Elemente etwas rasch hinweggegangen, kann aber bemerken, dass es sich hierbei lediglich um Bildung von Mittelwerten handelt. Sollte sich aber jemand finden, der sich um Einzelheiten interessiert, so bin ich zu Aufschlüssen gerne bereit.