

Vergrößerungsfaktor zum relativen Fehler bei Untergruppen in Stichproben

Autor(en): **Hülßen, Ellen / Cipriani, Josef**

Objektyp: **Article**

Zeitschrift: **Mitteilungen / Vereinigung Schweizerischer Versicherungsmathematiker = Bulletin / Association des Actuaire Suisses = Bulletin / Association of Swiss Actuaries**

Band (Jahr): **70 (1970)**

PDF erstellt am: **07.07.2024**

Persistenter Link: <https://doi.org/10.5169/seals-967035>

Nutzungsbedingungen

Die ETH-Bibliothek ist Anbieterin der digitalisierten Zeitschriften. Sie besitzt keine Urheberrechte an den Inhalten der Zeitschriften. Die Rechte liegen in der Regel bei den Herausgebern.

Die auf der Plattform e-periodica veröffentlichten Dokumente stehen für nicht-kommerzielle Zwecke in Lehre und Forschung sowie für die private Nutzung frei zur Verfügung. Einzelne Dateien oder Ausdrucke aus diesem Angebot können zusammen mit diesen Nutzungsbedingungen und den korrekten Herkunftsbezeichnungen weitergegeben werden.

Das Veröffentlichen von Bildern in Print- und Online-Publikationen ist nur mit vorheriger Genehmigung der Rechteinhaber erlaubt. Die systematische Speicherung von Teilen des elektronischen Angebots auf anderen Servern bedarf ebenfalls des schriftlichen Einverständnisses der Rechteinhaber.

Haftungsausschluss

Alle Angaben erfolgen ohne Gewähr für Vollständigkeit oder Richtigkeit. Es wird keine Haftung übernommen für Schäden durch die Verwendung von Informationen aus diesem Online-Angebot oder durch das Fehlen von Informationen. Dies gilt auch für Inhalte Dritter, die über dieses Angebot zugänglich sind.

Vergrößerungsfaktor zum relativen Fehler bei Untergruppen in Stichproben

Von Ellen Hülsen und Josef Cipriani

1. Problemstellung

Bei den klassischen Anwendungen des Stichprobenverfahrens, z. B. in der Biologie, sind die gesuchten Masszahlen sehr oft die arithmetischen Mittel, die dann auf ihre Abweichung entweder vom entsprechenden Mittelwert der Grundgesamtheit oder von Schätzwerten aus anderen Stichproben geprüft werden. Das Hauptgewicht liegt auf der Anordnung von Versuchsreihen und der Beurteilung der Abweichung zwischen verschiedenen Versuchsreihen bzw. Stichproben.

Die Problemstellung für die Stichprobenverfahren in der amtlichen Statistik ist in der Regel ganz anders. Gegeben ist ein bestimmtes Tabellenprogramm, gesucht sind die Summenwerte (Totalwerte) in den einzelnen Tabellenfeldern. Die Stichprobe bietet gegenüber der Vollerhebung eine Kostenersparnis. Allerdings muss der Benutzer der Statistik unter Umständen auf gewisse Merkmalskombinationen verzichten, deren Auszählung bei einer Vollerhebung durchaus möglich ist. Bei einer Stichprobe aber wären diese Ergebnisse zu ungenau.

Die Planung einer Stichprobenerhebung beginnt damit, das gewünschte Tabellenprogramm und die Erhebungseinheit festzulegen. Als zweites stellt man sich die Frage, welche Unterlagen als Grundgesamtheit zur Verfügung stehen, und wählt dementsprechend die Stichprobeneinheit. Als nächstes folgt die Festlegung des Stichprobenumfanges. Dazu sollte sich der Bearbeiter Rechenschaft geben können über die zu erwartenden Stichprobenfehler und die Möglichkeiten einer Schichtung der Grundgesamtheit. In diesem Stadium der Planung stellt sich oft ein besonderes Problem. Stimmt nämlich die Stichprobeneinheit nicht mit der Erhebungseinheit überein, dann stimmt auch die Gliederung der Tabellen nicht mit der Schichtung der Grundgesamtheit überein. In den einzelnen Tabellenfeldern hängen daher bei quantitativen Merkmalen

sowohl die Merkmalswerte als auch die Zahl der Merkmalsträger vom Zufall ab. Die Bildung von Tabellenfeldern oder *Untergruppen* vergrößert in diesem Falle den auf der Variabilität der Merkmalswerte beruhenden Stichprobenfehler um den sogenannten *Untergruppeneffekt*. Ein Untergruppeneffekt tritt z. B. auch dann ein, wenn zwar Stichprobeneinheit und Erhebungseinheit übereinstimmen, die Schichtung der Grundgesamtheit aber nur nach *einem* Merkmal durchgeführt wird, und die übrigen Gliederungsmerkmale der Tabellen aus der Stichprobe geschätzt werden müssen.

Es ist das Ziel dieser Arbeit, in den Anhangtabellen Grundlagen für die Schätzung der Untergruppeneffekte bei der Planung und Auswertung von Stichprobenerhebungen bereitzustellen.

2. Wirkung des Untergruppeneffektes

Es sei n die Anzahl Elemente in der Stichprobe aus der Grundgesamtheit von N Elementen. Von diesen n Einheiten gehören nur n_g zur betrachteten Untergruppe g mit dem gewünschten Merkmal. Mit $p = \frac{n_g}{n}$ werde der Anteil der Untergruppe g an der gesamten Stichprobe

bezeichnet, mit q dessen Komplement $1 - p = \frac{n - n_g}{n}$. Bei freier

Hochrechnung lautet dann die Schätzfunktion x'_g für den Totalwert x_g

$$x'_g = \frac{N}{n} x_g = Np \frac{x_g}{n_g} = Np \bar{x}_g \quad (1)$$

mit

$$x_g = \sum_{i=1}^{i=n_g} x_{ig}.$$

Die geschätzte Fehlervarianz von x'_g ist gegeben durch

$$\begin{aligned} s_{x'_g}^2 &= \varphi N \left(\frac{n}{n-1} pq \bar{x}_g^2 + \frac{n_g-1}{n-1} s_g^2 \right) \\ &\approx \varphi N (pq \bar{x}_g^2 + ps_g^2), \end{aligned} \quad (2)$$

worin $s_g^2 = \frac{1}{n_g - 1} \sum (x_{ig} - \bar{x}_g)^2$ die Streuung der Einzelwerte der Gruppe g in der Stichprobe

und $\varphi = \frac{1-f}{f} = \frac{N-n}{n}$

den sich auf Grund des Auswahlgesetzes $f = \frac{n}{N}$ ergebenden Auswahlfaktor bedeutet.

Aus den Gleichungen (1) und (2) ergibt sich die geschätzte relative Fehlervarianz von x'_g :

$$\begin{aligned} v_{x'_g}^2 &= \frac{s_{x'_g}^2}{x'_g} = \frac{1}{N^2 p^2 \bar{x}_g^2} \varphi N (pq \bar{x}_g^2 + p s_g^2) \\ &= \frac{\varphi}{N} \left(\frac{q}{p} + \frac{s_g^2 / \bar{x}_g^2}{p} \right) \\ &= \frac{\varphi}{N} \left(\frac{q}{p} + \frac{V^2}{p} \right), \end{aligned} \tag{3}$$

wenn man für den Variationskoeffizienten der Untergruppe V schreibt.

Gleichung (3) kann auch geschrieben werden:

$$v_{x'_g}^2 = v_{n'_g}^2 + \frac{\varphi}{N} \frac{V^2}{p} \tag{4}$$

n'_g ist Schätzfunktion für die Anzahl N_g der Einheiten der Untergruppe g in der Grundgesamtheit.

Der erste Term rechts in Gleichung (4) entspricht der geschätzten relativen Fehlervarianz von n'_g und zeigt damit den Einfluss des Zufalls bei der Auswahl der Einheiten, während sich der zweite Term auf die Streuung der Merkmalswerte bezieht.

Ist $p = 1$, so fällt wegen $q = 0$ der erste Term fort, es liegt der Normalfall ohne Untergruppen vor mit der relativen Fehlervarianz

$$v_{x'}^2 = \frac{\varphi}{N} V^2. \tag{5}$$

Der Quotient

$$A^2 = \frac{v_{x_g}^2}{v_x^2} = \frac{q + V^2}{V^2 p}$$

gibt ein Mass für die Vergrösserung der relativen *Varianz* durch den Untergruppeneffekt, in Abhängigkeit von Anteil p und Variationskoeffizient V . Für praktische Zwecke geeigneter ist jedoch die Quadratwurzel aus dieser Grösse,

$$A = \frac{1}{V} \sqrt{\frac{1}{p} (q + V^2)}, \quad (6)$$

welche nun direkt den Vergrösserungsfaktor zum relativen *Fehler* darstellt.

3. Anhangtabelle

Die Tabelle gibt den Vergrösserungsfaktor A gemäss Gleichung (6) zu Variationskoeffizienten von 0,05 bis 2,00 in Schritten von 0,05.

Der Anteilswert p in der Vorspalte läuft dabei von 0,005 bis 1,000, und zwar von 2,5% bis 97,5% in Schritten von 2,5%, während die Schrittlänge am Anfang und am Ende auf 0,5% verkürzt ist.

Als zweite Zeile ist im Tabellenkopf das Quadrat der darüber stehenden Variationskoeffizienten aufgeführt, da man in der Praxis manchmal diese Grösse bevorzugt. (Die Tabellenwerte wurden auf einer Tischrechenmaschine Olivetti Programma 101 vollautomatisch errechnet).

4. Literatur

Stichproben in der amtlichen Statistik. Statistisches Bundesamt, Wiesbaden (W. Kohlhammer GmbH, Stuttgart).

Hans Kellerer, Eine Verallgemeinerung des einfachen Urnenmodells und ihre Anwendung in der Stichprobentheorie (Allgemeines Statistisches Archiv, 39. Bd. 1955).

Zusammenfassung

Es wird ein spezielles, bei der Anwendung von Stichproben in der amtlichen Statistik jedoch nicht selten anzutreffendes Problem behandelt. Die beigefügte Tabelle gestattet dem Praktiker, sich rasch darüber zu orientieren, mit welcher Vergrößerung des relativen Fehlers durch den «Untergruppeneffekt» er zu rechnen hat.

Résumé

La présente étude a pour objet un problème spécial que l'on rencontre toutefois assez souvent lorsque l'on procède à des sondages dans les statistiques officielles. Le tableau ci-joint permet au statisticien de connaître rapidement l'augmentation de l'erreur relative que provoque l'«effet de sous-groupes».

