

Stichproben-Fehler

Die Verbraucher-Analyse basiert auf einer repräsentativen Zufallsstichprobe, die in Anlage und Durchführung der Europäischen Norm DIN/EN 15707:2008 „Printmedienanalysen - Begriffe und Dienstleistungsanforderungen“ entspricht. Dabei müssen die zufallsbedingten Abweichungen („Stichproben-Fehler“) berücksichtigt werden, die sich über mathematisch-statistische Verfahren berechnen lassen.

Der Stichproben-Fehler wird wie folgt geschätzt:

$$e = t * \sqrt{2} * \sqrt{\frac{p * (100 - p)}{n}}$$

- e = Stichproben-Fehler in Prozent
- t = Sicherheitsgrad (bei 95%: $t = 1,96$)
- p = Stichprobenanteil des Merkmals in Prozent
- n = (Teil-) Stichprobenumfang

In dieser Formel wird aufgrund der mehrstufigen Auswahl der in diesem Fall geklumperten Flächenstichprobe ein Designfaktor von $\sqrt{2}$ berücksichtigt, um den Stichproben-Fehler nicht zu unterschätzen.

Die Schätzung des relativen Merkmalsanteils (P) in der Grundgesamtheit (in Prozent) ergibt sich aus:

$$P = p \pm e$$

Der folgenden Tabelle der Schwankungsbreiten wurde eine Sicherheitswahrscheinlichkeit von 95% zugrundegelegt, wie in der Sozialforschung üblich.

Mit abnehmender Fallzahl werden die Ergebnisse immer „zufälliger“: So kann das tatsächliche Ergebnis beispielsweise auf Basis von 300 Fällen bei einem Produkt, das in dieser Stichprobe 5% der Befragten besitzen bzw. konsumieren, in der Grundgesamtheit mit einer Sicherheit von 95% zwischen 1,5% und 8,5% liegen. Das entspricht einem Schwankungsbereich von $\pm 3,5\%$. Während bei einer Fallzahl von 400 der Schwankungsbereich noch $\pm 3,0\%$ beträgt und bei steigender Fallzahl immer kleiner wird.

Der Problematik zu geringer Basiszahlen soll durch den Hinweis Rechnung getragen werden, dass **Zählungen mit Basis-Fallzahlen unter 300 für die Titelauswahl unbrauchbar** sind.

Tabelle der Schwankungsbreiten

$P = p \pm t * \sqrt{2} * \sqrt{\frac{p * (100 - p)}{n}}$		Sicherheitsgrad 95,0% = 1,96 σ $t = 1,96$															P = Anteil des Merkmals in der Grundgesamtheit $p =$ Stichprobenanteil des Merkmals $n =$ Stichprobenumfang								
		300	400	500	600	700	800	900	1000	1100	1200	1300	1400	1500	1600	1700	1800	1900	2000	2500	3000	4000	6000	8000	n
5% ±	3,5	3,0	2,7	2,5	2,3	2,1	2,0	1,9	1,8	1,7	1,7	1,6	1,6	1,6	1,5	1,5	1,4	1,4	1,4	1,2	1,1	1,0	0,8	0,7	5% ±
10% ±	4,8	4,2	3,7	3,4	3,1	2,9	2,8	2,6	2,5	2,4	2,3	2,2	2,1	2,1	2,0	2,0	2,0	1,9	1,9	1,7	1,5	1,3	1,1	0,9	10% ±
15% ±	5,7	4,9	4,4	4,0	3,7	3,5	3,3	3,1	3,0	2,9	2,7	2,6	2,6	2,5	2,4	2,3	2,3	2,3	2,2	2,0	1,8	1,6	1,3	1,1	15% ±
20% ±	6,4	5,5	5,0	4,5	4,2	3,9	3,7	3,5	3,3	3,2	3,1	3,0	2,9	2,8	2,7	2,6	2,5	2,5	2,2	2,0	1,8	1,4	1,2	1,0	20% ±
25% ±	6,9	6,0	5,4	4,9	4,5	4,2	4,0	3,8	3,6	3,5	3,3	3,2	3,1	3,0	2,9	2,8	2,8	2,7	2,4	2,2	1,9	1,5	1,3	1,1	25% ±
30% ±	7,3	6,4	5,7	5,2	4,8	4,5	4,2	4,0	3,8	3,7	3,5	3,4	3,3	3,2	3,1	3,0	3,0	2,9	2,8	2,5	2,3	2,0	1,6	1,4	30% ±
35% ±	7,6	6,6	5,9	5,4	5,0	4,7	4,4	4,2	4,0	3,8	3,7	3,5	3,4	3,3	3,2	3,1	3,0	3,0	2,8	2,6	2,4	2,1	1,7	1,5	35% ±
40% ±	7,8	6,8	6,1	5,5	5,1	4,8	4,5	4,3	4,1	3,9	3,8	3,6	3,5	3,4	3,3	3,2	3,1	3,0	2,7	2,5	2,2	1,8	1,5	1,3	40% ±
45% ±	8,0	6,9	6,2	5,6	5,2	4,9	4,6	4,4	4,2	4,0	3,8	3,7	3,6	3,4	3,3	3,3	3,3	3,2	3,1	2,8	2,5	2,2	1,8	1,5	45% ±
50% ±	8,0	6,9	6,2	5,7	5,2	4,9	4,6	4,4	4,2	4,0	3,8	3,7	3,6	3,5	3,4	3,3	3,3	3,2	3,1	2,8	2,5	2,2	1,8	1,5	50% ±
55% ±	8,0	6,9	6,2	5,6	5,2	4,9	4,6	4,4	4,2	4,0	3,8	3,7	3,6	3,4	3,3	3,3	3,3	3,2	3,1	2,8	2,5	2,2	1,8	1,5	55% ±
60% ±	7,8	6,8	6,1	5,5	5,1	4,8	4,5	4,3	4,1	3,9	3,8	3,6	3,5	3,4	3,3	3,2	3,1	3,0	2,7	2,5	2,2	1,8	1,5	1,3	60% ±
65% ±	7,6	6,6	5,9	5,4	5,0	4,7	4,4	4,2	4,0	3,8	3,7	3,5	3,4	3,3	3,2	3,1	3,0	3,0	2,6	2,4	2,1	1,7	1,5	1,3	65% ±
70% ±	7,3	6,4	5,7	5,2	4,8	4,5	4,2	4,0	3,8	3,7	3,5	3,4	3,3	3,2	3,1	3,0	2,9	2,8	2,5	2,3	2,0	1,6	1,4	1,2	70% ±
75% ±	6,9	6,0	5,4	4,9	4,5	4,2	4,0	3,8	3,6	3,5	3,3	3,2	3,1	3,0	2,9	2,8	2,8	2,7	2,4	2,2	1,9	1,5	1,3	1,1	75% ±
80% ±	6,4	5,5	5,0	4,5	4,2	3,9	3,7	3,5	3,3	3,2	3,1	3,0	2,9	2,8	2,7	2,6	2,5	2,5	2,2	2,0	1,8	1,4	1,2	1,0	80% ±
85% ±	5,7	4,9	4,4	4,0	3,7	3,5	3,3	3,1	3,0	2,9	2,7	2,6	2,6	2,5	2,4	2,3	2,3	2,2	2,0	1,8	1,6	1,3	1,1	0,9	85% ±
90% ±	4,8	4,2	3,7	3,4	3,1	2,9	2,8	2,6	2,5	2,4	2,3	2,2	2,1	2,1	2,0	2,0	1,9	1,9	1,7	1,5	1,3	1,1	0,9	0,7	90% ±
95% ±	3,5	3,0	2,7	2,5	2,3	2,1	2,0	1,9	1,8	1,7	1,7	1,6	1,6	1,5	1,5	1,4	1,4	1,4	1,2	1,1	1,0	0,8	0,7	0,5	95% ±

Leser-Beispiel: In einer Zielgruppe (=Unterstichprobe) vom Umfang $n = 1000$ besitzen bzw. konsumieren 30% der Befragten ein Produkt A. Aus der Tabelle entnimmt man die Fehlergrenzen $\pm 4,0\%$. Mit einer Wahrscheinlichkeit von 95% liegt dann der wahre Anteil in der Grundgesamtheit bei $30\% \pm 4,0\%$, d.h. zwischen 26,0% und 34,0%.