

Grundgesamtheit:

Die Bevölkerung eines Landes wird einer Volkszählung oder einer Wahl unterzogen. Alle Bürger werden erfaßt. Eine Fülle von Daten liegt vor.

Für diese Daten ist die beschreibende Statistik ungeheuer wichtig – schließlich muss man auswerten, zusammenfassen, darstellen.

Die beurteilende Statistik dagegen ist hier überflüssig. Die angefallenen Daten sind nicht zufällig, sondern sicher. Rechnet man etwa mit Hilfe der beschreibenden Statistik das Durchschnittsalter aller Einwohner aus, so stimmt das. Es gibt keinen Zweifel.

Stichprobe:

Wird hingegen, wegen des ungeheueren Aufwandes statt der Volkszählung nur eine gewisse Anzahl von Personen in eine Untersuchung einbezogen, so sind die Ergebnisse dieser Befragung keinesfalls sicher übertragbar auf die Gesamtbevölkerung.

Denn nun hat der Zufall seine Hand im Spiel – zufällig anders ausgewählte Personengruppen könnten durchaus andere Aussagen liefern.

Nehmen wir eine Maschine, die eine Million Plastikteller für Campingfreunde herstellt. Wenn jeder, wirklich jeder einzelne Teller hinsichtlich seiner Qualität untersucht wird, dann ist kein Zweifel am Ergebnis möglich. Die Aussage „3% Ausschuss“ ist wahr und damit sicher.

Wird aber, weil ökonomischer, nur eine gewisse Anzahl von Tellern zufällig entnommen und untersucht, dann ist die Aussage „3% Ausschuss“ nicht unkritisch verallgemeinerbar.

Diese zufällig erstellte Anzahl aus einer Gesamtheit nennt man Stichprobe.

Eine Stichprobe wird gezogen, wenn der Aufwand für die Untersuchung der Grundgesamtheit zu hoch ist – oder, wenn es schlicht unmöglich ist, die Grundgesamtheit zu analysieren. Um Glühbirnen z.B. auf ihre Lebensdauer zu testen, müßte man alle zerstören.

Normalverteilung:

Die Formulierung „...die Stichproben stammen aus normalverteilten Grundgesamtheiten...“ ist so zu verstehen, dass die Gesamtheit gleichartiger Einheiten, die hinsichtlich eines bestimmten Merkmals X untersucht werden, normalverteilt (Glockenkurve) ist. Dies ist keine Eigenschaft der Merkmalsträger, sondern der interessierenden Merkmalswerte, d.h. der Zufallsgröße X .

Eine aus der Grundgesamtheit nach dem Zufallsprinzip herausgegriffene Teilmenge mit n Elementen wird als Zufallsstichprobe vom Umfang n bezeichnet. Die Auswahl der Elemente muss

a) wahllos

b) unabhängig

voneinander erfolgen und jedes Element der Grundgesamtheit muss die gleiche Chance haben gezogen zu werden.

Die Formulierung „...man berechne den Stichprobenmittelwert m ...“ ist so zu verstehen, dass nicht von der Menge der Personen der „Mittelwert“ berechnet wird (wie auch?) sondern genauer:

Von jedem ausgewählten Stichprobenelement (Person) wird der Merkmalwert (Körpergröße)

beobachtet. Die beobachteten Merkmalwerte x_1, x_2, \dots, x_n sind dann Realisierungen der

Zufallsvariablen X und heißen Stichprobenwerte. Wenn also der „Mittelwert einer Stichprobe“ zu berechnen ist, dann ist also der „Mittelwert der Stichprobenwerte“ gemeint.