Summen – Zusammenfassung

Stellt euch eine Stichprobe vor wie einen Schrank voller Schubladen, die jeweils nur eine Sache beinhalten können. Jede Schublade hat einen Namen: die erste Schublade heißt X_1, die 2. Schublade heißt X_2 usw. Wenn also der Schrank ‚n' Schubladen hat, dann heißt die letzte X_n.

Nehmen wir jetzt an, dass wir einen Schrank mit 5 Schubladen haben. In jeder Schublade gibt es eine Statistik-Note. Nehmen wir an, dass in jeder Schublade eine 3 ist. Somit haben wir einen Schrank mit folgendem Inhalt:

$$\mathbf{X_1} = 3,\ \mathbf{X_2} = 3,\ \mathbf{X_3} = 3,\ \mathbf{X_4} = 3,\ \mathbf{X_5} = 3$$

Lasst uns alle diese 3en addieren:

$$\mathbf{X_1} + \mathbf{X_2} + \mathbf{X_3} + \mathbf{X_4} + \mathbf{X_5} = 3 + 3 + 3 + 3 + 3 = 15$$

Dies kann auch geschrieben werden mir dem $\sum$ Symbol:

$$\mathbf{X_1} + \mathbf{X_2} + \mathbf{X_3} + \mathbf{X_4} + \mathbf{X_5} = \sum_{i=1}^{5} X_i$$

wobei X_i=3 für i=1,2,3,4 und 5

$$\mathbf{X_1} + \mathbf{X_2} + \mathbf{X_3} + \mathbf{X_4} + \mathbf{X_5} = \sum_{i=1}^{5} \mathbf{3} = 15$$

Merkt, dass wenn alle X_i Schubladen den gleichen Wert haben und der Wert im Summensymbol steht, dann wird sich nie auf **i** bezogen. Wenn das passiert, ist das Resultat dieser Summe einfach der Wert (in diesen Beispiel ‚3') multipliziert mit der Anzahl der Schulbladen (in diesen Beispiel ‚5'); Also:

$$\mathbf{X_1} + \mathbf{X_2} + \mathbf{X_3} + \mathbf{X_4} + \mathbf{X_5} = \sum_{i=1}^{5} \mathbf{3} = 3*5 = 15$$

Falls wir nicht 5 sondern 245 Schubladen hätten, die eine 3 drin haben, dann wäre die Summe aller dieser Schubladen:

$$\mathbf{X_1} + \mathbf{X_2} \ldots\ldots + \mathbf{X_{245}} = \sum_{i=1}^{245} \mathbf{3} = 3*245 = 735$$

Allgemein gilt, wenn man eine Zahl ‚**Zahl**' ‚**n**'-mal addiert, dann wird diese Summe als $\sum_{i=1}^{n} \boldsymbol{Zahl}$ bezeichnet und ihre Resultat ist:

$$\sum_{i=1}^{n} Zahl = n \times Zahl \quad \mathbf{(1)}$$

Nehmen wir jetzt an, dass unser Schrank mit 5 Schubladen folgende Noten enthält:

$$\mathbf{X_1} = 1,\ \mathbf{X_2} = 2.3,\ \mathbf{X_3} = 4,\ \mathbf{X_4} = 3.3,\ \mathbf{X_5} = 1.7$$

Jetzt wollen wir alle diese Noten addieren:

$$\mathbf{X_1} + \mathbf{X_2} + \mathbf{X_3} + \mathbf{X_4} + \mathbf{X_5} = 1 + 2.3 + 4 + 3.3 + 1.7 = 12.3$$

Dies ist mit der Summe Symbol $\sum$ kurzer geschrieben:

$$\mathbf{X_1} + \mathbf{X_2} + \mathbf{X_3} + \mathbf{X_4} + \mathbf{X_5} = \sum_{i=1}^{5} X_i = 12.3$$

Was ist, wenn wir nicht die originalen Werte addieren wollen, sondern die quadrierte Werte? Das sieht dann so aus:

$$1^2 + 2.3^2 + 4^2 + 3.3^2 + 1.7^2 = 36.07$$

Die Kurze Form wäre:

$$1^2 + 2.3^2 + 4^2 + 3.3^2 + 1.7^2 = X_1^2 + X_2^2 + X_3^2 + X_4^2 + X_5^2 = \sum_{i=1}^{5} X_i^2$$

$$\sum_{i=1}^{5} X_i^2 = 36.07$$

Falls jede Schublade den gleichen Wert beinhaltet, z.B. ‚3', dann benutzen wir (1):

$$\sum_{i=1}^{5} 3^2 = 5 \times 3^2 = 45$$

Eigenschaften von Summen:

1. Das Resultat einer Summe ist eine Zahl. Zum Beispiel $\overline{X} = \frac{\sum_{i=1}^{n} x_i}{n}$ (der Mittelwert) ist eine Zahl.
2. Wenn ‚a' eine Zahl ist, dann $\sum_{i=1}^{n} a = n \times a$
3. Wenn ‚a' eine Zahl ist, dann $\sum_{i=1}^{n} aX_i = a \sum_{i=1}^{n} X_i$
4. Nehmen wir an, dass wir die Inhalte 2 gleich großer Schränke X und Y (gleiche Anzahl Schubladen) addieren wollen, dann $\sum_{i=1}^{n} (X_i + Y_i) = \sum_{i=1}^{n} X_i + \sum_{i=1}^{n} Y_i$

Anwendung der Summen Eigenschaften: Die Stichprobe Varianz

Die Varianz s^2 wird so berechnet:

$$s^2 = \frac{\sum_{i=1}^{n}(x_i - \overline{x})^2}{n-1}$$

Die Summe $\sum_{i=1}^{n}(x_i - \overline{x})^2$ wird als SAQ_x bezeichnet. Wir werden die Eigenschaften von Summen benutzen, um SAQ_x umzuformen:

Binomische Formel:

$$\sum_{i=1}^{n}(x_i - \overline{x})^2 = \sum_{i=1}^{n}(x_i^2 - 2\overline{x}x_i + \overline{x}^2)$$

Benutzen wir Eigenschaft 4: $\sum_{i=1}^{n}(X_i + Y_i) = \sum_{i=1}^{n}X_i + \sum_{i=1}^{n}Y_i$:

$$\sum_{i=1}^{n}(x_i - \overline{x})^2 = \sum_{i=1}^{n}x_i^2 + \sum_{i=1}^{n}(-2\overline{x}x_i) + \sum_{i=1}^{n}\overline{x}^2$$

Für den zweiten Term, wobei $-2\overline{x}$ eine Zahl ist (Eigenschaft 1), die mit X_i multipliziert wird, können wir Eigenschaft 3 benutzen:

$$\sum_{i=1}^{n}(-2\overline{x}x_i) = -2\overline{x}\sum_{i=1}^{n}x_i$$

$\overline{x}^2$ in dem dritten Term ist auch eine Zahl (Eigenschaft 1), also benutzen wir Eigenschaft 2:

$$\sum_{i=1}^{n}\overline{x}^2 = n \times \overline{x}^2$$

Also ist SAQ_x jetzt:

$$\sum_{i=1}^{n}(x_i - \overline{x})^2 = \sum_{i=1}^{n}x_i^2 - 2\overline{x}\sum_{i=1}^{n}x_i + n\overline{x}^2$$

Im zweiten Term können wir mit ‚n' multiplizieren und durch ‚n' dividieren und entsprechend einordnen:

$$\sum_{i=1}^{n}(x_i - \overline{x})^2 = \sum_{i=1}^{n}x_i^2 - 2n\overline{x}\frac{\sum_{i=1}^{n}x_i}{n} + n\overline{x}^2$$

Aber $\frac{\sum_{i=1}^{n} x_i}{n} = \bar{x}$, also:

$$\sum_{i=1}^{n}(x_i - \bar{x})^2 = \sum_{i=1}^{n} x_i^2 - 2n\bar{x}^2 + n\bar{x}^2$$

Und

$$\sum_{i=1}^{n}(x_i - \bar{x})^2 = \sum_{i=1}^{n} x_i^2 - n\bar{x}^2 - n\bar{x}^2 + n\bar{x}^2$$

Letztendlich:

$$SAQ_x = \sum_{i=1}^{n}(x_i - \bar{x})^2 = \sum_{i=1}^{n} x_i^2 - n\bar{x}^2 \quad (2)$$

Für zu Hause:

Die Kovarianz ist $s_{xy} = \frac{\sum_{i=1}^{n}(x_i - \bar{x})(y_i - \bar{y})}{n-1}$, wobei der Term $\sum_{i=1}^{n}(x_i - \bar{x})(y_i - \bar{y}) = SAQ_{xy}$ heißt.

SAQ_{xy} ist auch: $SAQ_{xy} = \sum_{i=1}^{n}(x_i - \bar{x})(y_i - \bar{y}) = \sum_{i=1}^{n} x_i y_i - n\bar{x}\bar{y}$; Versuch mal, dies zu überprüfen mit der Eigenschaften von Summen.

Häufig gemachte Fehler:

1. $SAQ_x = \sum_{i=1}^{n}(x_i - \bar{x})^2 = \sum_{i=1}^{n} x_i^2$ ist falsch!

Korrekt ist: $SAQ_x = \sum_{i=1}^{n}(x_i - \bar{x})^2 = \sum_{i=1}^{n} x_i^2 - n\bar{x}^2$

2. $SAQ_{xy} = \sum_{i=1}^{n}(x_i - \bar{x})(y_i - \bar{y}) = \sum_{i=1}^{n} x_i y_i$ ist falsch!

Korrekt ist $SAQ_{xy} = \sum_{i=1}^{n}(x_i - \bar{x})(y_i - \bar{y}) = \sum_{i=1}^{n} x_i y_i - n\bar{x}\bar{y}$

3. $SAQ_{xy} = \sum_{i=1}^{n}(x_i - \bar{x})(y_i - \bar{y}) = (x_i - \bar{x})(y_i - \bar{y})$ ist falsch! (Das Summensymbol kann nicht weggemacht werden.

4. $\sum_{i=1}^{n} x_i y_i = \sum_{i=1}^{n} x_i \sum_{i=1}^{n} y_i$ ist falsch! (Das funktioniert mit der Summe 2 Schränke - Eigenschaft 4, aber nicht mit der Multiplikation.)