Beispiel zur einfachen VA für unabhängige Stichproben

FRAGESTELLUNG

Die Wirkung des neuen Getreide-Düngemittels „Nutri 3000“ (N) soll mit dem bisher gebräuchlichen Mittel „Plantomax“ (P) verglichen werden. Von 28 Weizenfeldern (in vergleichbarer Lage) werden hierzu je 10 mit N und mit P gedüngt. Die restlichen 8 bleiben (quasi als „Kontrollgruppe“) ohne Dünger (O).

Anmerkung: In der Praxis würde man vermutlich auch unter Kontrollbedingung 10 Felder testen. Für dieses fiktive Beispiel wurden bewusst unterschiedliche Stichprobengrößen gewählt, um zu zeigen, dass diese für die VA nicht gleich sein müssen.

Die Erträge, gemessen in t/ha (Tonnen pro Hektar), waren wie folgt:

	N	P	O
	7.2	6.3	5.8
	7.7	7.1	6.4
	6.9	5.9	7.0
	7.1	6.5	5.9
	8.2	6.8	6.5
	6.4	6.1	6.1
	7.0	5.9	6.8
	6.7	6.5	7.1
	7.8	7.0	
	6.9	6.6	
$\sum_i x_{ij}$:	71.9	64.7	51.6
n_j:	10	10	8
$\bar{x}_{\cdot j}$:	7.19	6.47	6.45
$\sum_i x_{ij}^2$:	519.69	420.23	334.52
s_j^2:	0.303	0.180	0.243

Anzahl der Messwerte $\ldots N = 28$

Anzahl der Stichproben $\ldots k = 3$

Prüfen der Voraussetzungen

1.) Intervallskalierung der Daten

Die gemessene Dimension, Masse pro Fläche (hier in der Einheit t/ha angegeben), ist sogar rationalskaliert. Diese Voraussetzung ist also erfüllt.

2.) Normalverteilung in jeder Stichprobe

Die Prüfung der Normalverteiltheit von Stichproben wird erst später durchgenommen. Hier wird ohne Überprüfung angenommen, dass die Voraussetzung erfüllt ist.

3.) Homogenität der Varianzen

Da die Stichproben etwa gleich groß sind, kann der *Cochran*-Test verwendet werden:

$$C = \frac{s^2_{max}}{\sum_j s^2_j} \quad df = n^* - 1$$

$$(mit\ n^* = n\ von\ s^2_{max})$$

$$C = \frac{0.303}{0.303 + 0.180 + 0.243} \approx 0.417 \quad df = 9$$

Der kritische Wert für $\alpha = 0.05$, $df = 9$ und $k = 3$ ist:

$$C_{krit.} = 0.617$$

In diesem Fall gilt:

$$C = 0.417 < 0.617 = C_{krit.}$$

Ergebnis nicht signifikant $\Rightarrow$ Varianzen homogen. $\surd$

Aufstellen der Hypothesen

$H_0 \ldots \mu_N = \mu_P = \mu_O$

$H_1 \ldots \mu_N \neq \mu_P$ oder $\mu_N \neq \mu_O$ oder $\mu_P \neq \mu_O$

Berechnen der Varianzschätzungen

Wir müssen folgende Werte ermitteln:

$$\begin{aligned}
\sum_{ij} x_{ij}^2 &= \sum_i x_{i1}^2 + \sum_i x_{i2}^2 + \sum_i x_{i3}^2 = \\
&= 519.69 + 420.23 + 334.52 = 1274.44
\end{aligned}$$

$$\sum_{ij} x_{ij} = 71.9 + 64.7 + 51.6 = 188.2$$

$$\left(\sum_{ij} x_{ij}\right)^2 = 188.2^2 = 35419.24$$

$$QT = \sum_{ij} x_{ij}^2 - \frac{(\sum_{ij} x_{ij})^2}{N} = 1274.44 - \frac{35419.24}{28} \approx 9.467$$

$$\begin{aligned}
QZ &= \sum_j \frac{(\sum_i x_{ij})^2}{n_j} - \frac{(\sum_{ij} x_{ij})^2}{N} = \\
&= \left(\frac{71.9^2}{10} + \frac{64.7^2}{10} + \frac{51.6^2}{8}\right) - \frac{35419.24}{28} \approx 3.417
\end{aligned}$$

$$QI = QT - QZ = 9.467 - 3.417 = 6.05$$

$$\hat{\sigma}_1^2 = \frac{QZ}{k-1} \approx 1.709$$

$$\hat{\sigma}_0^2 = \frac{QI}{N-k} = 0.242$$

Quelle	QS	df	$\hat{\sigma}^2$
total	$QT = 9.467$	27	
zwischen	$QZ = 3.417$	2	$\hat{\sigma}_1^2 = 1.709$
innerhalb	$QI = 6.05$	25	$\hat{\sigma}_0^2 = 0.242$

Prüfgröße

$$F = \frac{\hat{\sigma}_1^2}{\hat{\sigma}_0^2} = \frac{1.709}{0.242} \approx 7.06$$

$$mit \;\; df_1 = 2, \;\; df_2 = 25$$

Vergleich mit dem kritischen Wert

Für $df_1 = 2$, $df_2 = 25$ und $\alpha = 0.05$ ist der kritische Wert bei einseitiger Fragestellung ($\hat{\sigma}_1^2 < \hat{\sigma}_0^2$ wird von vornherein ausgeschlossen):

$$F_{krit.} = 3.39$$

Das bedeutet, dass $F = 7.06 > 3.39 = F_{krit.}$ signifikant ist. Also ist die H_0 zu verwerfen, und es gilt (laut H_1):

$$\mu_N \neq \mu_P \text{ oder } \mu_N \neq \mu_O \text{ oder } \mu_P \neq \mu_O$$

Interpretation

Um festzustellen, welcher der drei möglichen Fälle, die die Alternativhypothese einschließt, gilt, wäre ein sog. *Post-Hoc*-Test durchzuführen. Da dies den Rahmen dieses Lernpfads übersteigt, müssen wir in den Daten nach Anhaltspunkten für eine Interpretation suchen. Am besten eignen sich hierfür üblicherweise die Stichprobenmittelwerte:

Der durchschnittliche Ertrag mit dem Düngemittel Plantomax beträgt 6.47 t/ha, fällt also nur geringfügig größer als bei ungedüngten Feldern (6.45 t/ha) aus. Die Wirkung ist also kaum spürbar.

Das neue Mittel Nutri 3000 bewirkt hingegen einen deutlich gesteigerten Ernteertrag von durchschnittlich 7.19 t/ha.

Das Ergebnis der VA ist somit, dass das neue Mittel signifikant besser ist als das bisher verwendete.