

Versuchsdaten sagen alleine wenig aus

Warum braucht es Statistik im Versuchswesen?

„Traue keiner Statistik, die du nicht selbst gefälscht hast!“ ist ein bekannter Spruch und verkörpert in gewissem Maße das Misstrauen vieler gegenüber der Statistik. Die als theoretisch empfundene Disziplin ist jedoch in Zeiten großer Datenmengen und komplexer Fragestellungen wichtiger denn je, auch in der Landwirtschaft. Die Bereitstellung von Daten allein ist noch nicht ausreichend, erst durch eine adäquate Statistik kann aus Daten Wissen generiert werden.



Ohne Wiederholungen innerhalb eines Feldes lässt sich der Einfluss von Bodenunterschieden oder anderen Einflussfaktoren nicht quantifizieren. Foto: Becker

Eine Kernaufgabe der Statistik ist es, mit Hilfe der vorhandenen Daten Entscheidungsträgern die Arbeit zu erleichtern, zum Beispiel den Landwirten bei der Sortenentscheidung zu unterstützen. Im Folgenden sollen Grundzüge und Begrifflichkeiten aus der Statistik im landwirtschaftlichen Versuchswesen verständlich erläutert werden.

Statistik als Werkzeug zur leichteren Entscheidungsfindung

Häufig werden hinter dem Begriff Statistik komplexe Vorgänge verstanden, die einen hohen Grad an mathematischem Verständnis voraussetzen. Dabei beginnt Statistik bereits bei der Beobachtung und Datenerhebung. Die Antwort auf die Frage, wie eine einfache Definition von Statistik aussehe, beantwortet ChatGPT folgendermaßen: „Statistik ist die Wissenschaft, die sich damit beschäftigt, Daten zu sammeln, zu ordnen, auszuwerten und zu interpretieren, um Informationen zu gewinnen und Entscheidungen zu unterstützen.“

zen. Kurz gesagt: Statistik hilft uns, aus Zahlen Sinn zu machen.“

Im Fall des pflanzenbaulichen Versuchswesens begleitet die Statistik einen Versuch von der Planung bis zur Auswertung. Ohne statistische Planung und Auswertung wären die erhobenen Daten wertlos. Sie müssen in einen Zusammenhang gebracht und interpretiert werden, um daraus Informationen zu erhalten, die dann verarbeitet und verknüpft werden müssen, um daraus Wissen zu schaffen. Erst dieses Wissen ermöglicht Handeln und Entscheidungen. Diese Schritte von Daten bis zum Wissen werden stets durch die Statistik begleitet. Um also Wissen zu generieren, brauchen wir den „Werkzeugkasten“ der Statistik.

Zu Beginn jedes Versuches steht eine Frage, die es zu beantworten gilt. Fragestellungen im landwirtschaftlichen Versuchswesen kommen typischerweise aus der Produktion, wie zum Beispiel „Welche Sorte bringt in welcher Region den höchsten Ertrag?“ Entsprechend werden neue Züchtungen an verschie-

denen Versuchsstandorten unter verschiedenen Umweltbedingungen getestet, um Aussagen über die Leistungsfähigkeit einer Sorte treffen zu können.

Dabei liegt das Hauptaugenmerk oft auf dem Ertrag, den es zu steigern gilt. Der Ertrag wird durch viele Faktoren beeinflusst, wie beispielsweise die Verfügbarkeit an Wasser und Nährstoffen, Sonneneinstrahlung und vieles mehr. Der Einfluss dieser Faktoren kommt jedoch nicht in jedem Jahr und an jedem Standort gleichermaßen zum Tragen.

Die Bewertung der Sortenleistung kann daher nur in einem systematischen Ansatz beantwortet werden, der sich der Statistik bedient. Erst dadurch wird eine solide Entscheidungsgrundlage geschaffen.

Anwendung von Statistik am Beispiel der Sortenbeurteilung

Exaktversuche werden stets mit Wiederholungen durchgeführt. In den Veröffentlichungen zu Ergebnissen von Sortenversuchen findet man oft die Begriffe Grenzdifferenz und Varianz. Anhand eines Beispiels soll verdeutlicht werden, welche Bedeutung diese Begriffe haben und in welcher Beziehung sie zueinanderstehen. Die Grafik visualisiert diese Beziehungen und zeigt den schematischen Ablauf der Schritte eines Versuches.

Ein Landwirt möchte herausfinden, welche Sorte für seinen Betrieb die Beste ist. Auf einem Schlag werden drei für ihn in Frage kommende Sorten in jeweils einem Streifen angebaut. Pflanzenschutz, Düngung und sonstige Pflegemaßnahmen werden für alle Sorten konstant gehalten. Sorte A erzielt 90 dt/ha, Sorte B 85 dt/ha und Sorte C 92 dt/ha. Rein nach den Zahlen wäre es klar, Sorte C wäre zu bevorzugen.

Allerdings fehlen Informationen, um eine Entscheidung zu treffen. Die bessere Leistung von Sorte C könnte auch darauf zurückzuführen sein, dass sie zufällig in einem Teil des Schlages mit günstigeren Bodeneigenschaften stand. Es ist also nicht möglich zu schlussfolgern, welchen Einfluss die Sorte beziehungsweise andere Faktoren wie der Boden hatten. Die Datenlage reicht somit nicht aus, um ausschließen zu können, dass die Beobachtungen rein zufällig waren.

Zufall oder nicht?

Um den Zufall ausschließen zu können, muss der Einfluss anderer Faktoren berücksichtigt werden. Das Beispiel wird erweitert: Der Landwirt möchte nach wie vor wissen, wie der Faktor Sor-

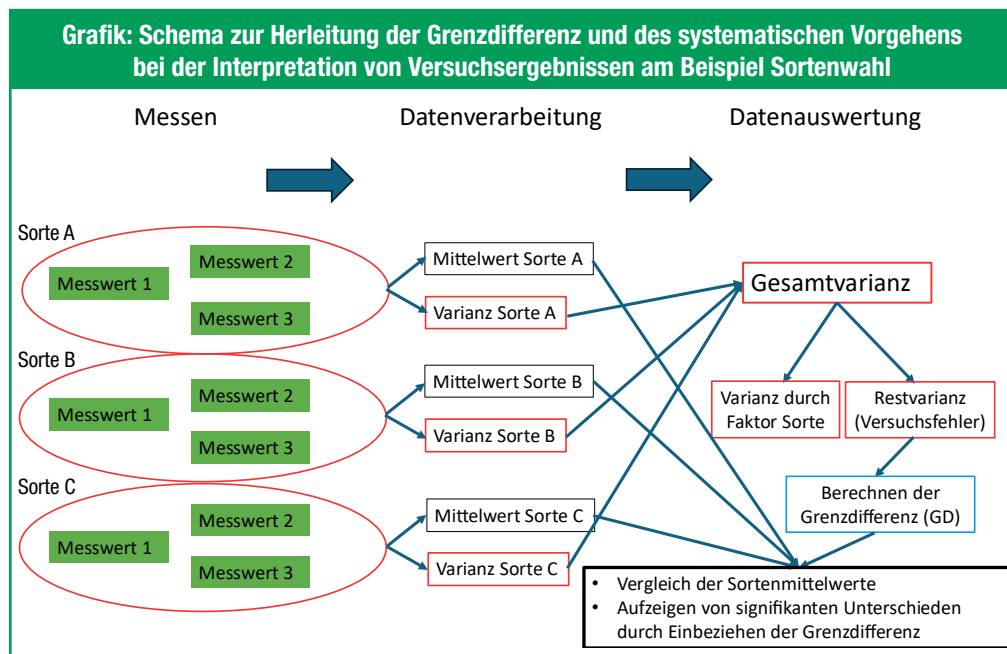
te den standortspezifischen Ertrag beeinflusst. Er legt für jede Sorte drei Streifen an, die er zufällig auf seinem Feld verteilt, das heißt das Feld wird in neun Streifen aufgeteilt. Er wiederholt den Anbau jeder Sorte, um zu schauen, ob die Sorten an einer anderen Stelle des Ackers ein ähnliches Ergebnis erzielen. Die Erntemengen werden für jeden Streifen separat erfasst.

Zum besseren Verständnis zeigt Tabelle 1 zwei Beispiele auf, welche die Bedeutung der Varianz verdeutlichen. In beiden Beispielen haben die Sorten A, B und C dieselben Mittelwerte erzielt. Der große Unterschied liegt in der Streuung der Daten, die durch die Varianz beschrieben wird. Im Beispiel 2 wird ersichtlich, dass die Erträge der Sorten zwischen den Streifen sehr starke Unterschiede aufweisen. Das gibt Anlass, sich Gedanken über Bodenunterschiede zu machen, die offensichtlich einen großen Einfluss auf den Ertrag haben. Erst durch die Wiederholung an unterschiedlichen Stellen eines Feldes kann dies festgestellt werden.

Ohne Wiederholungen innerhalb eines Feldes lässt sich der Einfluss von Bodenunterschieden oder anderen Faktoren nicht quantifizieren. Im Beispiel 1 sind die Unterschiede zwischen den Streifen nur marginal, dementsprechend ist die Varianz auch kleiner.

Beschränkung auf den Faktor Sorte

Wie soll der Landwirt sich nun entscheiden? Bei beiden Beispielen wäre bei einer alleinigen Betrachtung der Mittelwerte nach wie vor die Sorte C zu bevorzugen. Um aber die Sortenunterschiede nicht falsch zu interpretieren, muss die Varianz einbezogen werden. Welche Faktoren die Streuung verursacht haben, lässt sich zunächst nicht erkennen. Bodenunterschiede, Fehler in



der Versuchsdurchführung, Beschattungen von Teilflächen, das sind alles Faktoren, die die Varianz mitverursacht haben können. Die Varianz muss erst in verschiedene Teilvarianzen aufgeteilt werden. Da die Sorte in dem Beispiel im Fokus steht, möchten wir aus der Gesamtvarianz die durch den Faktor Sorte bedingte Varianz kennen. Einfacher ausgedrückt: Welchen Anteil an der Streuung des Ertrags verursacht die Sorte, welchen alle anderen Faktoren?

Den Anteil aller anderen Faktoren, die in einem Versuch nicht betrachtet werden, wird in der Statistik auch als „Versuchsfehler“ bezeichnet. In einer theoretischen Welt, in der alle Faktoren bis auf die Sorte gleich wären, gäbe es keinen Versuchsfehler. Der Mehraufwand durch das Anlegen von Wiederholungen wäre nicht nötig und die Frage nach der besten Sorte schnell beantwortet. Der Aufwand muss in einer multifaktoriellen Welt mit vielen äußeren

Einflüssen betrieben werden, um gesicherteres Wissen, zum Beispiel über die Sortenwahl, zu erhalten.

Sind Sortenunterschiede signifikant?

Nach Aufteilung der Varianz in ihre Einzelteile (Sorte / alle anderen Faktoren) kann die sogenannte Grenzdifferenz (GD) berechnet werden. Sie ist das statistische Maß, welches einen direkten Vergleich der Sortenmittelwerte und Aussagen über die Signifikanz ermöglicht. Im Beispiel 1 liegt die Grenzdifferenz bei 4,34, das heißt ist der Ertragsunterschied der Sortenmittelwerte größer als 4,34, ist die Sorte mit dem höheren Ertrag signifikant besser als die andere Sorte. Folgende Rechnungen werden also benötigt:

90 dt/ha (Sorte A) – 85 dt/ha (Sorte B) = 5 dt/ha, folglich ist Sorte A bei einer Grenzdifferenz von 4,34 dt/ha signifikant besser als Sorte B

92 dt/ha (Sorte C) – 90 dt/ha (Sorte A) = 2 dt/ha, folglich ist Sorte C nicht signifikant besser als Sorte A, die Differenz zwischen Sorte C und A kann nicht mit 95-prozentiger Sicherheit durch den Faktor Sorte erklärt werden

92 dt/ha (Sorte C) – 85 dt/ha (Sorte B) = 7 dt/ha, Sorte C ist signifikant besser als Sorte B.

Nach dieser Rechnung kann der Landwirt also die Sorte B für seinen Standort ausschließen, sie brachte signifikant weniger Ertrag als die beiden anderen Sorten. In Beispiel 2, bei der die Grenzdifferenz mit 36,63 aufgrund der hohen Varianz deutlich höher liegt als im Beispiel 1, ist keine Sorte signifikant besser beziehungsweise schlechter als die andere, da keine Sorte einen

Tabelle 1: Kornertrag in dt/ha im Versuch des Landwirtes mit unterschiedlicher Varianz und Grenzdifferenz (GD)

Kornertrag in dt/ha	Beispiel 1					
	Streifen 1	Streifen 2	Streifen 3	Mittelwert	Varianz	GD bei 5 % Irrtumswahrscheinlichkeit
Sorte A	90	89	91	90	1	
Sorte B	83	87	85	85	4	
Sorte C	94	92	90	92	4	
Beispiel 2*						
Sorte A	80	90	100	90	100	36,63
Sorte B	74	89	92	85	93	
Sorte C	68	108	100	92	448	

*das Beispiel 2 zeigt, dass eine große Streuung der Einzelwerte um den Mittelwert (dieser ist jeweils gleich zu Beispiel 1) zu hohen Varianzen und Grenzdifferenzen führt; hier ist keine Sorte signifikant besser beziehungsweise schlechter als die andere, da keine Sorte einen Mehrertrag von 36,63 dt/ha erzielen konnte.

Tabelle 2: Auszug aus dem Winterraps LSV 2024 in reduzierter Variante

relativ zu Bezugsbasis BB	Status	Hybriden (H)	Prüfjahr	Friedberg	Fritzlar	Korbach	Mittel
Mittelwert BB (dt/ha)				47,1	49,2	48,1	48,1
GD 5 % (relativ)				7,0	13,4	10,7	
Ludger	BB	H	>3	88	86	86	87
Ambassador	BB	H	>3	109	101	97	102
Otello KWS EU	BB	H	>3	101	95	103	100
Daktari	BB	H	>3	104	110	108	107
LG Activus	BB	H	>3	94	107	98	100
Allesandro KWS EU	BB	H	>3	95	106	106	102
Scotch	BB	H	3	99	94	98	97
PT303		H	2	105	89	98	97
LG Adonis	BB	H	3	107	103	100	103
Picard	BB	H	3	103	99	104	102
PT299		H	2	93	89	88	90
PT302		H	2	103	90	98	97
LG Auckland		H	2	94	72	84	83
Hermann		H	2	100	97	100	99
Vespa		H	2	108	106	103	106
Humboldt		H	2	106	100	113	107
Archivar		H	2	101	94	101	99
KWS Ambos		H	2	103	93	93	96
LG Ambrosius		H	1	99	98	96	98

Mehrertrag von 36,63 dt/ha erzielen konnte. Dieses fiktive Beispiel des kleinen Sortenversuches soll verdeutlichen, dass die reinen Mittelwerte nicht ausreichen können, um fundierte Entscheidungen treffen zu können.

Landessortenversuche und die Grenzdifferenz

In der Auswertung der Ergebnisse von Landessortenversuchen (LSV) wird die Grenzdifferenz ebenfalls berechnet. Tabelle 2 zeigt einen Ausschnitt des

Auf einen Blick

- Ohne (echte) Wiederholungen kann der sogenannte „Versuchsfehler“ nicht festgestellt werden. Die Beobachtungen können rein zufällig sein und sind statistisch nicht abgesichert.
- Die Grenzdifferenz zeigt die Signifikanzen zwischen den Sorten auf und gibt Sicherheit bei der Entscheidungsfindung.
- Je größer die Streuung der Daten in einem Versuch ist, desto höher fällt die Grenzdifferenz aus.
- Die Grenzdifferenz ist indirekt ein Ausdruck für die durch Zufallseinflüsse hervorgerufene Variabilität.

Dr. Klauk

Winterraps-LSV aus der Saison 2023/2024. Die Bezugsbasis (BB) sind die Sorten, die drei Jahre oder länger geprüft wurden; sie bilden die Referenz für den relativen Vergleich aller Sorten. Der Mittelwert aller Sorten der Bezugsbasis bildet die 100 Prozent.

In Tabelle 2 wird am Standort Friedberg deutlich, dass die Sorte Ludger mit 88 Prozent einen signifikant geringeren Ertrag hatte als die Sorte Ambassador mit 109 Prozent, da die relative Grenzdifferenz mit 7 Prozent geringer ist als die Differenz von 21 Prozent zwischen den beiden Sorten. Es konnte also nachgewiesen werden, dass die Sorte Ambassador mit 95-prozentiger Wahrscheinlichkeit einen höheren Ertrag

liefert als die Sorte Ludger. LG Activus hatte in diesem Jahr einen relativen Ertrag von 94 Prozent, der Unterschied zwischen LG Activus und Ludger war damit nicht signifikant. Die Grenzdifferenz ist somit eine zentrale Information für die Leserschaft von LSV-Berichten.

Das am Beispiel des Sortenversuches erläuterte Vorgehen ist allgemein anwendbar und beschränkt sich nicht nur auf Sortenversuche. Prinzipiell sind alle Versuche, bei den verschiedene Prüfglieder miteinander verglichen werden sollen, nach diesem Vorgehen auswertbar.

*Dr. Benjamin Klauk,
Fachinformation Pflanzenbau,
Landesbetrieb Landwirtschaft Hessen*