

Berechnung einer ANOVA für Gruppenvergleiche in R

Kursbegleitendes Material

Inhaltsverzeichnis

01.01 Einführung in die ANOVA für Gruppenvergleiche	1
01.02 Bevor wir rechnen	5
1. Datenmanagement	5
2. Vorbereitung in R.....	6
2.1 Installation von Paketen	6
2.2 Daten aus Excel Datei einlesen.....	7
01.03 Deskriptive Analyse	10
1. Grobe Visualisierung der Daten.....	11
2. Ausreißer	14
02.01 Berechnung der ANOVA.....	16
1. Wie sind die Ergebnisse der ANOVA zu interpretieren?.....	17
2. Effektstärke der ANOVA	18
3. Berichten der ANOVA	19
02.02 Voraussetzungsprüfung	20
1. Testung auf Normalverteilung der Residuen	20
2. Varianzhomogenität.....	23
Kleiner Exkurs	24
03.01. Berechnung einer robusten ANOVA mit Bootstrapping.....	25
03.02 Fazit der ANOVA und weiteres Vorgehen.....	28
04.01 Post hoc Tests	29
1. Post hoc t-Tests - Normales Vorgehen, ohne robuste Tests	29
Option 1: Haupteffekt Schmerzmittel.....	30
Option 2: Haupteffekt Dosis	33
Option 3: Signifikante Zweifachinteraktion aus Schmerzmittel * Dosis	34
04.02 Post hoc Yuen-Tests - Robustes Vorgehen	37
Option 1: Haupteffekt Schmerzmittel.....	38
Option 2: Haupteffekt Dosis	40
Option 3: Signifikante Zweifachinteraktion aus Schmerzmittel * Dosis	44



04.03 Geplante Kontraste	54
1. <i>Geplante Kontraste - normales Vorgehen</i>	54
2. <i>Geplante Kontraste mit Bootstrapping - Robustes Vorgehen</i>	60
05.01 Datenvisualisierung	63

01.01 Einführung in die ANOVA für Gruppenvergleiche

Wozu benötigst du eine ANOVA für Gruppenvergleiche?

ANOVAs für Gruppenvergleiche verwendest du immer dann, wenn du *mehr als zwei Gruppen* bzgl. einer metrischen, also stetigen Variablen (wie Größe in cm, Gewicht in kg, Alter in Jahren) vergleichen möchtest.

Warum? Weil die ANOVA für Gruppenvergleiche super effektiv ist und dir eine Menge Arbeit und Probleme erspart! Außerdem erlaubt sie dir bei Bedarf auch mehrere Gruppenvariablen in Kombination zu testen - das schafft kein T-Test!

Ein einfaches Beispiel einer ANOVA mit einer Gruppenvariablen: Angenommen, du möchtest den Mittelwert dreier Therapiegruppen A, B und C miteinander vergleichen. Bei nur zwei Gruppen (A vs. B) bräuchtest du nur einen einzigen Test, und die Sache wäre abgehakt. Möchtest du aber drei Gruppen vergleichen benötigst du insgesamt drei Tests, um alle Gruppenkombinationen zu testen (A vs. B, A vs. C und B vs. C).

Verdammt viele Tests, oder? Und genau da liegt das Problem.

Je mehr Gruppen verglichen werden sollen, desto stärker wächst die Zahl der benötigten T-Tests. Das ist nicht nur aufwendig, sondern auch aus statistischer Sicht problematisch. Ein einzelner Test ist nämlich dann signifikant, wenn der p-Wert kleiner als .05 ist. Damit akzeptierst du eine Wahrscheinlichkeit von 5% einen Unterschied zwischen Gruppen als signifikant einzustufen, obwohl da garkeiner ist (alpha-Fehler).

Bei *drei* Tests liegt diese Wahrscheinlichkeit schon bei 14% ($(1-(0,95^3))*100 = 14.26$). Das ist viel? Genau! Und deshalb nennt man diesen steilen Anstieg auch die *Alpha-Fehler-Inflation*.

Und um genau dieses Problem zu vermeiden, verwenden wir anstatt der vielen einzelnen T-Tests *eine* ANOVA. Mit ihr senken wir das Risiko einer falsch positiven Entscheidung zurück auf 5%. Bingo!

Wie das geht? Die ANOVA sagt dir schon im ersten Rechenschritt, ob es Sinn macht all die vielen T-Tests überhaupt zu rechnen – oder nicht. Sie erteilt die dabei die "Erlaubnis" *ob* und *wo genau* du nach Gruppenunterschieden suchen darfst. Das ist besonders wichtig, wenn du nicht nur den Einfluss einer Gruppenvariablen testest, sondern mehrerer. Ist die ANOVA nicht signifikant, vermeidest du somit nicht nur eine Alpha-Fehler-Inflation, sondern sparst dir unheimlich viel Zeit, da du schon sehr früh weißt, ob es sich überhaupt lohnt, alle Gruppen miteinander zu vergleichen.

Und wie macht man das? Das möchte ich dir anhand eines Datenbeispiels mit zwei Gruppenvariablen genau erklären. Im Folgenden siehst du Auszüge aus einem Excel Datensatz mit insgesamt 60 Personen. Alle Patienten haben in unserem Beispiel entweder das Schmerzmittel Ibuprofen oder Paracetamol (Gruppenvariable 1) erhalten. Die Dosis war dabei entweder gering,

mittel oder hoch (Gruppenvariable 2). Zusätzlich wurde die empfundene Wirkdauer in Stunden erfasst. Jetzt möchtest du wissen, ob die verabreichte Dosis einen Einfluss auf die Wirkdauer hat und ob dieser Dosisseffekt - je nach Schmerzmittel - unterschiedlich ist.

	A	B	C	D
1	id	Schmerzmittel	Dosis	Wirkdauer
2	1	Paracetamol	gering	1.40
3	2	Paracetamol	gering	3.83
4	3	Paracetamol	gering	2.43
11	10	Paracetamol	gering	2.33
12	11	Paracetamol	mittel	5.50
13	12	Paracetamol	mittel	5.50
20	19	Paracetamol	mittel	6.27
21	20	Paracetamol	mittel	5.17
22	21	Paracetamol	hoch	7.87
30	29	Paracetamol	hoch	7.77
31	30	Paracetamol	hoch	9.83
32	31	Ibuprofen	gering	5.07
40	39	Ibuprofen	gering	5.50
41	40	Ibuprofen	gering	3.23
42	41	Ibuprofen	mittel	6.57
51	50	Ibuprofen	mittel	9.10
52	51	Ibuprofen	hoch	8.50
53	52	Ibuprofen	hoch	8.80

Es handelt es sich hier also um eine klassische Studie, in der der Einfluss eines Wirkstoffs (Ibuprofen vs. Paracetamol) und verschiedener Dosierungen (gering vs. mittel vs. hoch) auf die Wirkdauer untersucht werden soll.

Deshalb spricht man bei den Gruppenvariablen Schmerzmittel und Dosis auch von „unabhängigen Variablen“ (oder im Kontext der ANOVA von „Faktoren“) und bei der Variable Wirkdauer von der „abhängigen Variablen“. Wir Forscher manipulieren also die Ausprägungen der unabhängigen Variablen (bzw. Faktoren) und schauen, wie sich das auf die abhängige Variable auswirkt.

Du merkst schon: Hier handelt es sich ausschließlich um kategoriale Gruppenvariablen, sie bilden also feste Klassen, die keinen fließenden Übergang haben. Ein Patient bekommt also Paracetamol ODER Ibuprofen, dazwischen gibt es in dieser Studie nichts (also vergiss Paraprufen, daraus wird nichts). Diese Klassenbildung muss auch so sein, denn nur dann kannst du Gruppen bilden und die brauchst du, um eine ANOVA zu berechnen.

Da sich in jeder Gruppe verschiedene Personen befinden spricht man auch von einem Studiendesign für Gruppenvergleiche (oder Englisch: Between-Subject's Design oder auch Zwischensubjekt Design auf Deutsch). Da in unserem Beispiel zwei Faktoren mit zwei bzw. drei Faktorstufen vorliegen ist es ein 2*3 between-subject's Design bzw. ein 2*3 Zwischensubjekt Design.

So viele Fachwörter– aber wat mutt dat mutt! Und dafür hast du es jetzt hinter dir!

Du hast also Patienten, die entweder Ibuprofen ODER Paracetamol bekommen, und zwar in einer niedrigen, einer mittleren ODER einer hohen Dosis erhalten. Dies resultiert somit in 6 verschiedenen Gruppen. In jeder der sechs Gruppen befinden sich 10 Personen. Das sind eigentlich zu wenig pro Gruppe, aber wir halten es in unserem Beispiel der Übersichtlichkeit halber einfach:

		Dosis		
		gering	mittel	hoch
Schmerzmittel	Ibuprofen	n = 10	n = 10	n = 10
	Paracetamol	n = 10	n = 10	n = 10

Infobox: In diesem Studiendesign handelt es sich um ein Design ohne jegliche Messwiederholung. Wenn sich in allen Gruppen die gleichen Personen befinden (d. h. wenn die gleichen Personen sechsmal zu verschiedenen Bedingungen gemessen wurden), dann schau am besten in unser Tutorial zur ANOVA für Messwiederholungen. Wenn beispielsweise jeweils 10 Personen Ibuprofen oder Paracetamol erhalten haben und beide Gruppen dann dreimal mit jeweils anderer Dosis gemessen wurden, dann bist du bei der ANOVA für Gruppe x Messwiederholung richtig.

Wie also wertet man solche Daten am besten aus? Deine 10 Schritte zum Glück:

1. Vorbereitung und Datenmanagement: Damit es richtig losgehen kann muss dein Datensatz ähnlich aussehen wie meine Excel-Liste. Dabei ist es kein Problem, wenn du nur einen oder sogar drei Faktoren hast. Es ist auch kein Problem, wenn die einzelnen Faktoren mehr oder weniger Faktorstufen aufweisen. Unsere Anleitung funktioniert immer!

Außerdem befinden sich in deinem Datensatz wahrscheinlich noch ein Haufen anderer Variablen. Auch das ist kein Problem. Ich gehe aber davon aus, dass dein Datensatz im Grunde gut aufbereitet ist. Wie man das macht und was es alles zu beachten gibt, zeige ich dir übrigens ganz genau in unserem Basiskurs „R - Lern die Basics!“.

2. Vorbereitungen in R: Im nächsten Schritt musst du deine Daten in R einlesen. Auch das Installieren von Paketen, die dir das Leben leichter machen, machen wir ganz am Anfang.
3. Berechnung deskriptiver Statistiken: Zu Beginn ist es wichtig, dass du dich erst einmal mit deinen Daten vertraut machst. Das umfasst das Berechnen von Mittelwerten und Standardabweichungen pro Gruppe, das Prüfen auf Ausreißer und auch eine erste Visualisierung deiner Daten. Wir zeigen dir wie das geht.



4. Berechnung der ANOVA: Anschließend berechnen wir die ANOVA. Diese gibt dir Aufschluss darüber, welche Gruppenvergleiche im Anschluss berechnet werden dürfen. Natürlich zeige ich dir auch, wie man die ANOVA berichtet.
5. Voraussetzungsprüfung: Bevor es zu den Gruppenvergleichen geht, müssen wir einige Voraussetzungen für die ANOVA prüfen. Sind Annahmen verletzt solltest du ein robustes Verfahren in Betracht ziehen und auf die komplette Analyse von der ANOVA bis hin zu den Post hoc-Tests anwenden. (Übrigens: Dass wir die ANOVA berechnen haben, bevor wir deren Annahmen prüfen, hat rein praktische Gründe und ist ein interner Trick, der dir viel Arbeit erspart).
6. Ggf. Berechnung der robusten ANOVA: Hat die Voraussetzungsprüfung ergeben, dass deine Daten nicht den Anforderungen der ANOVA genügen, zeigen wir dir, wie du ganz einfach eine robuste ANOVA berechnest, mit der du völlig auf der sicheren Seite bist. Du musst dein Studiendesign dabei nicht ändern, weil du – wie sonst immer- auf andere Verfahren wie den Kruskal-Wallis-Test ausweichen musst.
7. Fazit Berechnung ANOVA: Anhand der (signifikanten) Haupt- und Interaktionseffekte entscheidest du anhand unserer Anleitung, wie du weiter vorgehst, denn je nach Effekt, darfst du im Folgenden nur bestimmte Gruppenvergleiche durchführen. Wir zeigen dir, wie das geht!
8. Gruppenvergleiche: Hier prüfst du deine Hypothesen durch das Vornehmen von direkten Gruppenvergleichen. Dafür gibt es verschiedene Verfahren, z. B. den Vergleich aller Gruppen durch post-hoc Tests oder das Vornehmen geplanter, also hypothesenbasierter Kontraste. Hast du dich für eine robuste ANOVA entschieden, solltest du auch hier auf robuste Tests zurückgreifen.
9. Egal welcher Weg für deine Daten der Beste ist: Wir zeigen dir wie du deine Resultate (im Ergebnisteil) und dein statistisches Vorgehen (im Methodenteil) in deiner wissenschaftlichen Arbeit korrekt berichtest – natürlich auf Deutsch und auf Englisch!
10. Datenvisualisierung: Nichts bringt deine Ergebnisse so gut rüber wie eine richtig schöne Grafik – wir zeigen dir wie du sie erstellst.

Fertig? Dann können wir loslegen!