

Regressionsanalyse - Seminar / Kurs von DoE Consulting

In diesem eintägigen Seminar „Regressionsanalyse“ erlernen Sie sowohl die Grundlagen der Regression als auch speziellere Gebiete der Regressionsanalyse, die üblicherweise nicht so bekannt aber trotzdem sehr hilfreich bei der Prozessoptimierung sind.

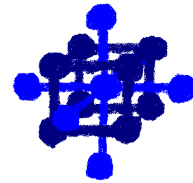
Die Grundlagen werden durch die **einfache lineare Regression** gelegt. Hier lernen Sie insbesondere die Korrelationsanalyse, die Konfidenz- und Prognoseintervalle und die Residualanalyse kennen.

Der zweite Teil dieses Seminars ist die **Multiple Regression**. Es werden Regressionsmodelle für mehr als zwei Faktoren erstellt und erläutert. Diese Modelle können Wechselwirkungsterme enthalten, um eine schiefe Antwortflächenebene oder einen Knick der Antwortfläche zu modellieren. Und quadratische Terme werden genutzt, um eine Krümmung abzubilden.

Im dritten Teil wird die **Anwendung dieser Response Surface Modelle** und die konkrete Ausführung aller notwendigen Berechnungsschritte ausführlich und anhand von konkreten Beispielen behandelt.

Im vierten Teil des Seminars wird die **Binäre Logistische Regression** behandelt. Diese muss verwendet werden, wenn die Zielgröße Y diskrete Werte und die Faktoren X_i stetige Werte annehmen. Beispielsweise könnte die Zielgröße sein „Pakete, die unpünktlich geliefert werden“ und ein Einflussfaktor wäre „Entfernung vom Ursprung zum Ziel der Lieferung“. Hier ist also die Entfernung ein stetiger Wert, der genau in Metern oder Kilometern gemessen werden kann. Die Zielgröße „unpünktliche Pakete“ würde als Anteil unpünktlicher Pakete von allen Paketen einer Stichprobe berechnet. Wir lernen, wie man dies mit Hilfe des sog. Wahrscheinlichkeitsverhältnisses (Odds ratio) modelliert. Die Notwendigkeit der Wahrscheinlichkeitsfunktion wird anschaulich erklärt und anschließend deren Anwendung in praxisnahen Beispielen erläutert.

Im fünften Teil werden weitere spezielle Themen erläutert. Die Gefahr der **Multi-Kolinearität** tritt häufig bei der Analyse historischer Daten auf. Wenn die signifikanten Faktoren korrelieren, können Fehler bei der Regressionsanalyse passieren. So werden falsche Faktoren ausgewählt oder die Faktoreffekte werden falsch berechnet und insgesamt werden damit schlechte Vorhersagegleichungen aufgestellt. Wie kann man dies entdecken und vermeiden?



Bei der Modellierung mit diskreten X_i ist die Verwendung von sog. **Indikatorvariablen** häufig sinnvoll und hilfreich. Die Einfügung solcher Indikatorvariablen in der Auswertung mit Minitab wird anhand eines Beispiels detailliert erläutert.

Mit Hilfe der **Residualanalyse** kann man erkennen, ob Terme in der Vorhersagegleichung fehlen. Wenn die Residuen nicht normalverteilt sind und die Auftragung der Residuen über die angepassten Modellwerte wie ein „V“ aussieht, erkennt man, dass also größere Residuen bei hohen und niedrigen Y-Werten im Vergleich zu mittelgroßen Y-Werten auftreten. Hier fehlt ein **Wechselwirkungsterm** in der Vorhersagegleichung!

Zeigt zudem auch die Auftragung der Residuen über die X-Werte eines Faktors z. B. X_1 diese V-Form, dann ist **Krümmung** übersehen worden und es sollte ein quadratischer Term X_1^2 im Modell aufgenommen werden.

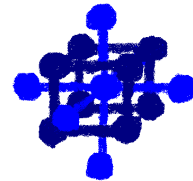
Diese Zusammenhänge werden mit anschaulichen Beispielen und ausführlichen Auswertungen erklärt.

Neben dem Auffinden optimaler Faktoreinstellungen ist auch die Bestimmung von Spezifikationsgrenzen für die Faktoreinstellungen bzw. deren Schwankung sehr wichtig. Hierzu werden **Toleranzbereiche für jeden Faktoren** mittels Regression und Vorhersageintervalle ermittelt. So stellt man sicher, dass das Optimum der Zielgröße stets erreicht wird, man aber zudem auch eine maximale Flexibilität in der Einstellung der Faktoren hat. Die praktische Vorgehensweise wird detailliert an einem praxisnahen Beispiel vorgeführt.

Abgeschlossen wird das Seminar mit zwei umfangreichen Übungen, die das Erlernete nochmal zur Anwendung bringen und festigen.

Inhalte:

- Einleitung und Motivation
 - Die Funktionsgleichung $Y=f(X_i)$.
 - Überblick Hypothesentests, Regression als HT für stetige Y- und stetige X-Werte, t-Verteilung, t-Wert, Null- und Alternativ-Hypothese, alpha-Wert, p-Wert.
- Einfache lineare Regression
 - Korrelationsanalyse: Hängt Y von X ab? Wenn Ja, wie stark? Pearson-Koeffizient r , Korrelationsmatrix, Bestimmtheitsmaß R^2 , Spearman-Koeffizient.

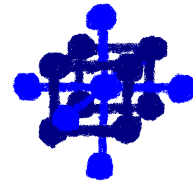


- Einfache lineare Regression: Darstellung der Anpassungslinie, Konfidenz- und Prognoseintervalle, p -, S -, R^2 -Werte.
- Residual-Analyse: Residuen und angepasste Werte, Prüfung auf Normalverteilung der Residuen, I-Chart, Histogramm und Auftragung der Residuen als Funktion der Anpassungen.
- Beispiel für eine quadratische Anpassung.

- Multiple Regression
 - Überblick: Wenn Y von mehreren X_i abhängt und die X_i korrelieren.
 - Geometrie und Modell: Von der Geraden zur Ebene (vom 2-Dimensionalen zum 3-Dimensionalen). Regressionsmodell für mehr als zwei X , mit Wechselwirkungstermen und mit Krümmung.
 - Multiple Regression als Response Surface Modell (Antwortflächenmodell) in Minitab.
 - Modellierung und Hypothesen Test: t -Test und ANOVA, Konfidenzintervalle.

- Response Surface Modelle (Antwortflächenmodelle)
 - Vorgehensweise bei der Auswertung und Modellerstellung.
 - Lack-of-fit (Modellfehler).
 - S und R^2 .
 - Reduzierung der Anzahl Modellparameter.
 - Graphische Darstellung der Vorhersagegleichung und Optimumsuche.
 - Response-Optimizer und überlagertes Konturdiagramm.

- Binäre Logistische Regression
 - Was ist BLR: Y diskrete/binäre Werte und X_i stetige Werte.
 - Wofür wird BLR benötigt: Wahrscheinlichkeit des „Erfolges“ vs. Ja/Nein.
 - Wie funktioniert BLR: Anpassung der Wahrscheinlichkeitsfunktion mit logarithmischer Funktion, anschauliche Betrachtung von Wahrscheinlichkeiten.
 - Wie macht man eine BLR: S -Kurve, Vorhersagegleichung, p -Werte, Wahrscheinlichkeitsverhältnis (Odds ratio) und Konfidenzintervall.
 - Wie interpretiert man die Ergebnisse.
 - Beispiele.
 - BLR mit mehreren X_i : Auswertung mit t -Test, Berechnung der Wahrscheinlichkeiten, Beispiel.
 - Binäre LR: 2 Niveaus für X_i , Ordinale LR: 3 und mehr Niveaus für X_i und X_i haben eine Ordnung (z. B. wenig/mittel/hoch), Nominale LR: wie OLR aber X_i haben keine Ordnung (z. B. Gelb/Schwarz/Rot).



- Spezielle Themen
 - Multi-Kolinearität = Korrelation der Faktoren und dessen Einfluss auf die Regression, Beispiel, Korrelationsanalyse, Reduzierung des Modells.
 - Modellierung mit diskreten X_i : Indikatorvariablen erstellen.
 - Wie erkennt man fehlende Terme: Residualanalyse.
 - Festlegung von Toleranzbereiche für X_i : Regression und Vorhersageintervalle.
- Übungen

Software:

Minitab wird als kostenlose und vollwertige 4-wöchige Testversion zur Verfügung gestellt.

Die Teilnehmer bringen ein **Notebook** mit (Admin-Rechte zur Installation der Statistik-Software Minitab notwendig).

Dokumentation:

Handout aller Powerpoint Folien in einem Ordner mit **250 Powerpoint Folien** auf 125 DIN-A4 Seiten. Diese enthalten die behandelten Theorien und deren Anwendung in Übungen mit vollständigen Lösungswegen.

Abschluss:

Teilnahmebescheinigung.

Schriftliche oder mündliche **Prüfung auf Anfrage** möglich.

Kursleitung:

Dipl.-Ing. **Dr.-Ing. Fischer**

Mehr als 20 Jahre Erfahrung in der Anwendung von Methoden der statistischen Optimierung. Speziell in Forschung, Entwicklung und Produktion der chemischen und pharmazeutischen Industrie sowohl als Angestellter als auch als Consultant.

Anzahl Teilnehmer je Kurs:

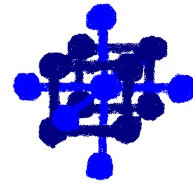
max. 6 Teilnehmer

Dauer:

1 Tag, 8 h, 9:00 bis 17:00 Uhr

Kursgebühr:

495,00 Euro zzgl. MwSt.



Aktuelle Termine:
siehe Homepage: www.doe-academy.de/seminare/

Inhouse Seminare nach Vereinbarung in Deutschland und der Schweiz
jederzeit möglich. Seminarinhalte nach Absprache variabel.

Auf Wunsch stehen wir auch nach dem Seminar mit zeitlich und inhaltlich
maßgeschneidertem **Consulting** zur Verfügung.

Anfragen und Buchung:
Dr.-Ing. Fischer
Tel. 0231 2000662
E-Mail: info@doe-consulting.de