

Modellierung der Preis-Absatz Funktion

Karsten Lübke

17.04.2018

1. Forschungsfrage

Das Preismanagement ist eine zentrale Frage für Unternehmen, aber die Abschätzung der Zusammenhänge ist nicht einfach, siehe z. B. (Steiner and Weber 2010).

Es existieren verschiedene Modellierungsansätze für Preis-Absatz Funktionen. Hier ist der Preis p die erklärende Variable, die Menge q die erklärte Variable: $q = f(p)$, z. B.:

- Linear: $q = a - b \cdot p$. Vgl. $y = \beta_0 + \beta_1 \cdot x$.
- Multiplikativ: $q = a \cdot p^{-b}$. Vgl. $y = \beta_0 \cdot x^{\beta_1} \Leftrightarrow \ln(y) = \ln(\beta_0) + \beta_1 \cdot \ln(x)$.

mit unbekannten Koeffizienten a, b (bzw. β_0, β_1), die mit Hilfe von Daten geschätzt werden können.

Welche der Modellierungsansätze kann Preis-Absatz Funktionen besser beschreiben?

2. Studiendesign

Die Zahlungsbereitschaft wurde erhoben anhand der Frage: "Würden Sie das Produkt zum Preis P kaufen?" vgl. z. B. (Adler 2003). Hier für eine fiktive organische Milch im Preisbereich $p = 0.19, 0.29, 0.39, \dots, 2.99$ (in Euro).

Die Teilnehmer*innen wurden anhand einer freiwilligen und anonymen Online Umfrage befragt.

3. Datenerhebung

Die Datenerhebung erfolgte über eine Verteilung des Umfragelinks im Freundes- und Bekanntenkreis. Die aufbereiteten Daten der Online Umfrage liegen als `csv` Datei vor. Es wurde je Preisstufe der Anteil der Personen ($n = 104$) berechnet, die das Produkt zum angegebenen Preis kaufen würden.

```
Milk <- read.csv2("http://www.statistix.org/Data/OrganicMilk.csv")
```

4. Datenanalyse

Datenbeschreibung:

```
# Ggfs. Paket mosaic installieren
# install.packages(mosaic)

# Paket laden
library("mosaic")

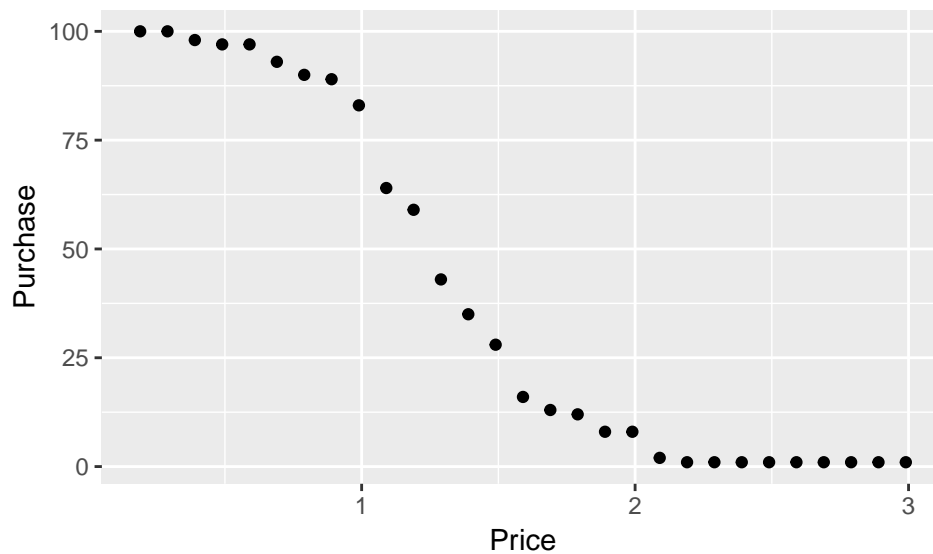
# Erste Datenübersicht
inspect(Milk)
```

```
##
## quantitative variables:
##      name  class min   Q1 median   Q3    max    mean      sd   n
## 1   Price numeric 0.19 0.89   1.59 2.29   2.99  1.59000  0.8514693 29
## 2 Purchase integer 1.00 1.00  16.00 89.00 100.00 39.44828 40.9691401 29
## missing
## 1      0
## 2      0
```

- Price: Preis der Milch.
- Purchase: Anteil der Personen, die die Milch zu dem Preis kaufen würden.

Grafische Analyse

```
gf_point(Purchase ~ Price, data = Milk)
```



Die grafische Analyse über das Streudiagramm zeigt einen fallenden Verlauf (je höher der Preis, desto geringer der Anteil der Leute, die das Produkt kaufen würden) als auch Sprungstellen.

Kennzahlen

Die Korrelationskoeffizienten nach Pearson und auch Spearman bestätigen den optischen Eindruck, wobei die Rangkorrelation (Spearman) größer ist:

```
cor(Purchase ~ Price, data = Milk, method = "pearson")
```

```
## [1] -0.9399529
```

```
cor(Purchase ~ Price, data = Milk, method = "spearman")
```

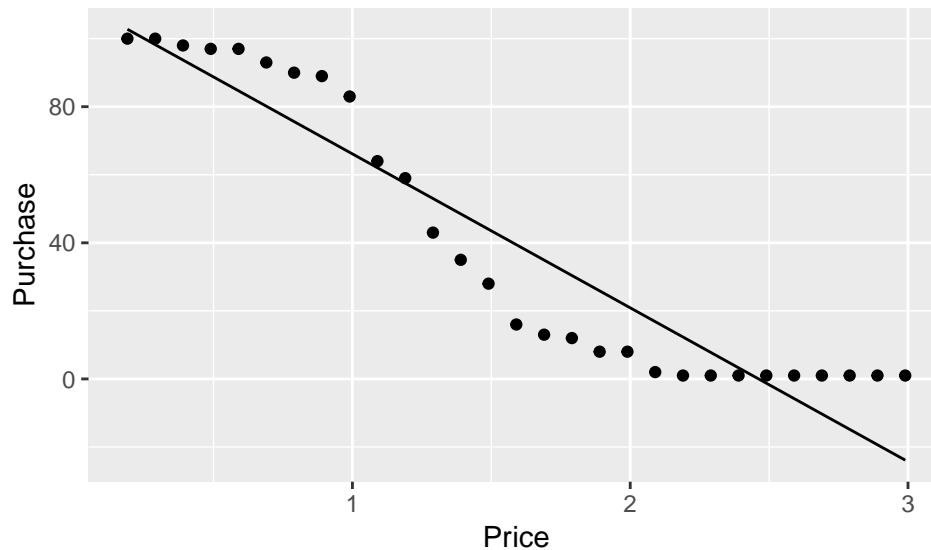
```
## [1] -0.9847357
```

5. Modellierung

Linear

Über den ganzen Wertebereich ist eine lineare Modellierung nicht sinnvoll:

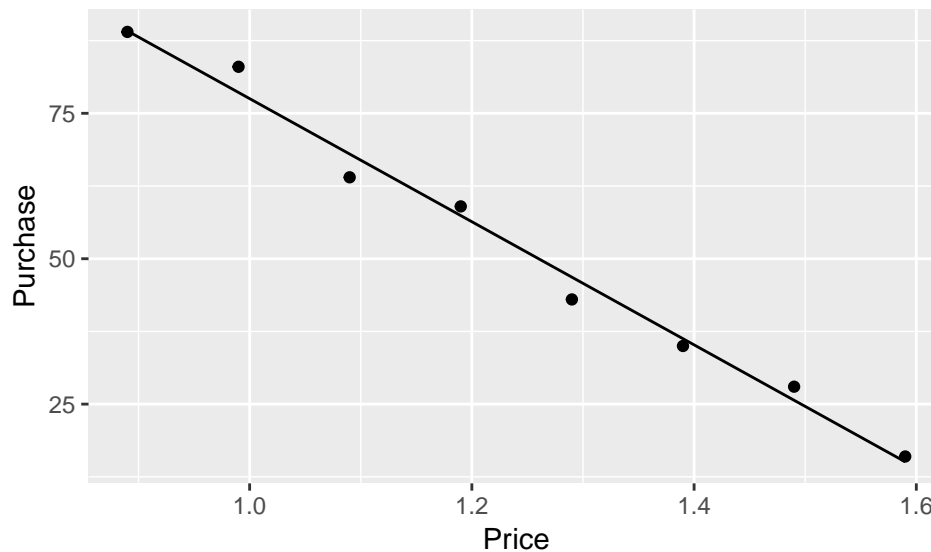
```
linear1 <- lm(Purchase ~ Price, data = Milk)
gf_point(Purchase ~ Price, data = Milk) %>%
  gf_line(fitted(linear1) ~ Milk$Price)
```



Eine lokale Anpassung für den Bereich von 0.89 bis 1.59:

```
Milk2 <- Milk %>%
  filter(0.89<=Price & Price<=1.59)

linear2 <- lm(Purchase ~ Price, data = Milk2)
gf_point(Purchase ~ Price, data = Milk2) %>%
  gf_line(fitted(linear2) ~ Milk2$Price)
```



```
summary(linear2)
```

```
##
## Call:
## lm(formula = Purchase ~ Price, data = Milk2)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -4.000 -1.896  0.375  1.771  4.417
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)  183.358      6.183   29.66 9.75e-08 ***
## Price       -105.833      4.903  -21.59 6.45e-07 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 3.178 on 6 degrees of freedom
## Multiple R-squared:  0.9873, Adjusted R-squared:  0.9852
## F-statistic: 465.9 on 1 and 6 DF,  p-value: 6.454e-07
```

Innerhalb des Wertebereiches zwischen 0.89€ und 1.59€ gilt als Schätzgleichung:

$$\hat{y} = 183.36 - 105.83 \cdot x$$

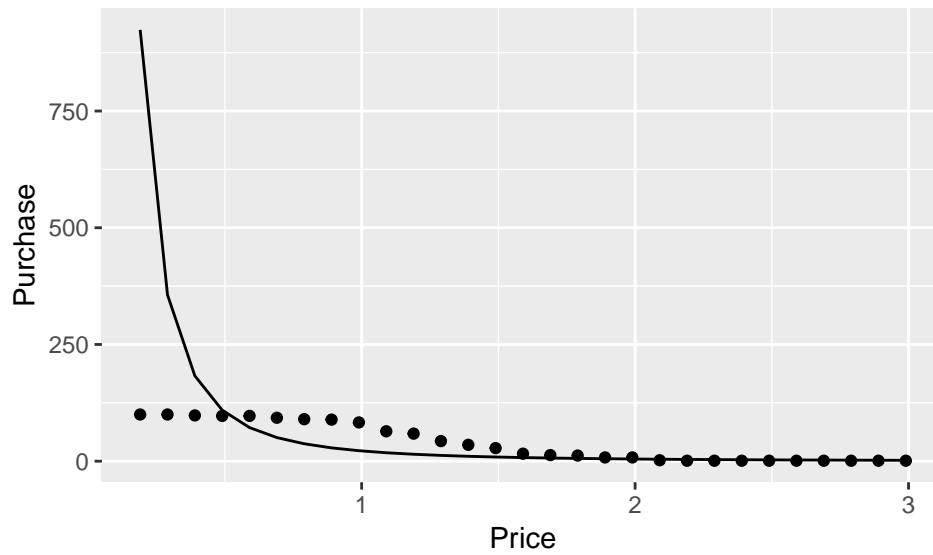
D. h., pro 0.10€ wird in diesem Wertebereich ein Rückgang des Mittelwertes der Zahlungsbereitschaft um 10.58% beobachtet. Eine Extrapolation sollte nicht erfolgen. Allerdings ist in diesem Bereich die Anpassung sehr gut: $R^2 = 0.987$.

Die Nullhypothese “Es gibt keinen linearen Zusammenhang zwischen Preis und Absatz”, d. h. $H_0 : \beta_P = 0$ wird mit einem P-Wert von 6.4541138×10^{-7} verworfen ($\alpha = 0.001$).

Multiplikativ

Über den ganzen Wertebereich ist auch eine multiplikative Modellierung nicht sinnvoll:

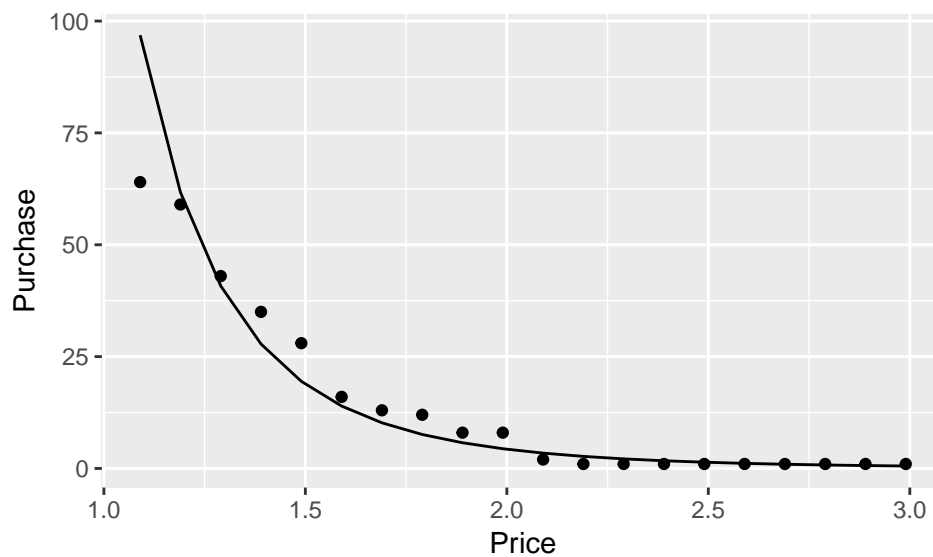
```
multi1 <- lm(log(Purchase) ~ log(Price), data = Milk)
gf_point(Purchase ~ Price, data = Milk) %>%
  gf_line(exp(fitted(multi1)) ~ Milk$Price)
```



Eine Anpassung im Wertebereich 1.09 bis 2.99 sieht wie folgt aus:

```
Milk2 <- Milk %>%
  filter(1.09<=Price & Price<=2.99)

multi2 <- lm(log(Purchase) ~ log(Price), data = Milk2)
gf_point(Purchase ~ Price, data = Milk2) %>%
  gf_line(exp(fitted(multi2)) ~ Milk2$Price)
```



```
summary(multi2)
```

```
##
## Call:
## lm(formula = log(Purchase) ~ log(Price), data = Milk2)
##
## Residuals:
```

	Min	1Q	Median	3Q	Max
##	-0.9922	-0.3535	0.1002	0.3392	0.6059

```
##
```

```
## Coefficients:
##           Estimate Std. Error t value Pr(>|t|)
## (Intercept)   5.0154     0.2580   19.44 1.57e-13 ***
## log(Price)   -5.1324     0.3514  -14.61 2.01e-11 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.4722 on 18 degrees of freedom
## Multiple R-squared:  0.9222, Adjusted R-squared:  0.9179
## F-statistic: 213.3 on 1 and 18 DF,  p-value: 2.012e-11
```

In diesem Bereich gilt in diesem Modell:

$$\hat{y} = e^{5.02} \cdot x^{-5.13}$$

D. h., die geschätzte Preiselastizität der Nachfrage liegt hier bei -5.13 . Auch hier ist die Anpassung gut: in diesem Bereich werden auf logarithmischer Skala $R^2 = 0.922$ der Variation der Nachfrage modelliert¹ und $H_0 : \beta_P = 0$ wird verworfen ($\alpha = 0.001$).

6. Schlussfolgerungen

In den vorliegenden Daten können weder das lineare noch das multiplikative Modell die Preis-Absatz Funktion über den ganzen Wertebereich modellieren - es sind nur lokale Anpassungen sinnvoll.

Außerdem wurde nur die fiktive Kaufabsicht einer Gelegenheitsstichprobe erhoben, es wurden keine realen Kaufdaten verwendet.

Ausblick

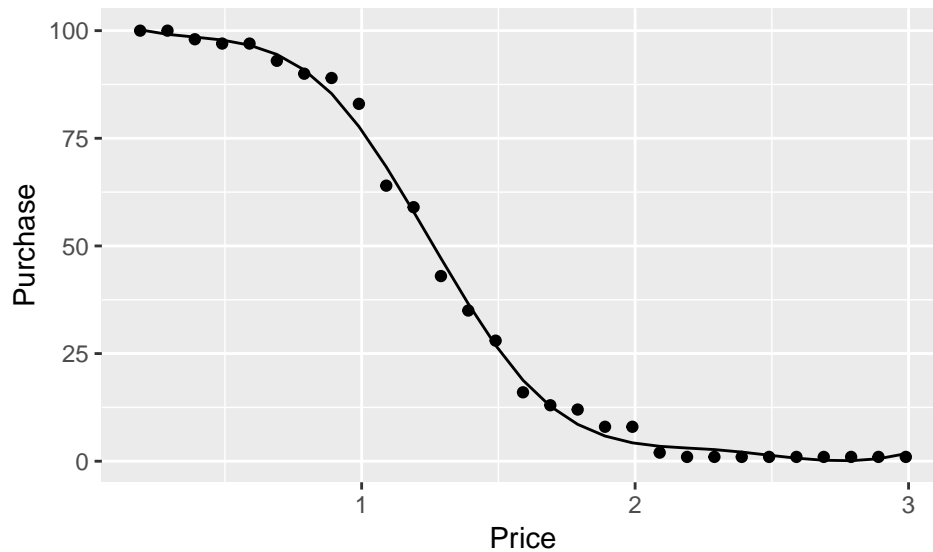
Eine flexible Alternative ist die Spline Regression. Hier wird der Wertebereich von x in Intervalle aufgeteilt und innerhalb dieser Intervalle werden Polynome angepasst, wobei die Übergänge geglättet werden, siehe z. B. (James et al. 2013).

```
# Paket laden
library(splines)

splines <- lm(Purchase ~ bs(Price,
                           knots=quantile(Price, probs=c(0.25, 0.50, 0.75)),
                           degree=3),
             data=Milk)

gf_point(Purchase ~ Price, data = Milk) %>%
  gf_line(fitted(splines) ~ Milk$Price)
```

¹ R^2 von transformierten und untransformierten Variablen y können nicht direkt verglichen werden.



Die Interpretation der Koeffizienten (vgl. `summary(splines)`) ist schwierig. Aber man kann erkennen, dass die Anpassung über den ganzen Wertebereich gut ist: $R^2 = 0.997$.

Danksagung

Die Daten wurden innerhalb der Bachelor Thesis zum Thema “Spline Regression as a Method to Estimate Price Response Functions” erhoben.

Anhang: Versionshinweise

- Datum erstellt: 2019-06-10
- R Version: 3.6.0

Verwendete Pakete:

- `mosaic` Version: 1.5.0
- `ggformula` Version: 0.9.1
- `splines` Version: 3.6.0

Literatur

Adler, Jost. 2003. “Möglichkeiten Der Messung von Zahlungsbereitschaften Der Nachfrager.” *Duisburger Arbeitspapiere Zum Marketing* 7.

James, Gareth, Daniela Witten, Trevor Hastie, and Robert Tibshirani. 2013. *An Introduction to Statistical Learning*. Springer.

Steiner, Winfried J, and Anett Weber. 2010. “Methoden Zur Schätzung von Preis-Absatz-Funktionen.” *Wirtschaftsstudium-WISU* 39 (1): 121–28.