Fachhochschule Nordwestschweiz
Hochschule für Wirtschaft

Kommentar zu "Kronthaler: Statistik angewandt"

Ein bisschen Mathematik muss schon sein!

Studie

Autor: Helmut Vetter
Ort, Datum: Arlesheim, 02.09.2014

Diese Arbeit wurde mit TexLive erstellt.

Kommentar zu "Kronthaler: Statistik angewandt"
Ein bisschen Mathematik muss schon sein!

Autor
Vetter, Helmut
Schillerweg 2
CH-4144 Arlesheim
061 599 51 09
helmut.vetter@fhnw.ch

Auftraggeberschaft
Fachhochschule für Wirtschaft
Tanner, Christian

Arlesheim, September 2014

Ehrenwörtliche Erklärung
Ich versichere, dass ich die vorliegende Arbeit selbstständig und ohne Benutzung anderer als der im Literaturverzeichnis angegebenen Quellen und Hilfsmittel angefertigt habe.

Die wörtlich oder inhaltlich den im Literaturverzeichnis aufgeführten Quellen und Hilfsmitteln entnommenen Stellen sind in der Arbeit als Zitat bzw. Paraphrase kenntlich gemacht.

Diese Arbeit ist noch nicht veröffentlicht worden. Sie ist somit weder anderen Interessenten zugänglich gemacht noch einer anderen Prüfungsbehörde vorgelegt worden.

Arlesheim, 02.09.2014

Helmut Vetter

Helmut Vetter

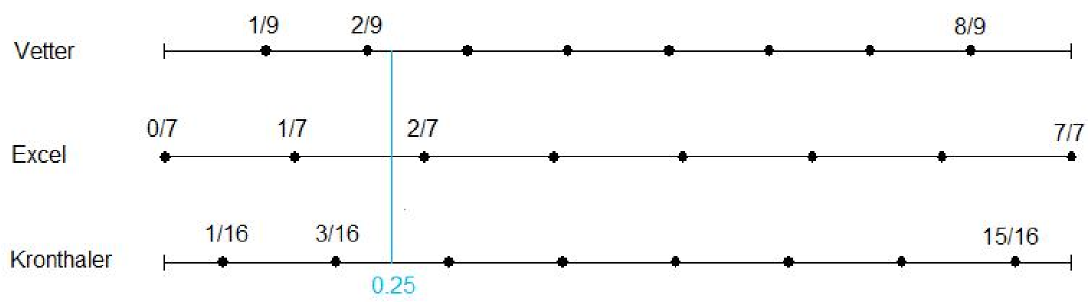
Management Summary

Auf Anregung von Christian Tanner für mehr Praxisbezug im Mathematikunterricht habe ich das Buch von Franz Kronthaler "Statistik angewandt" gelesen.
Vom Ansatz her ist das Buch ordentlich. In mathematischen Details enthält es einige Fehler.
Ich möchte hier 4 Punkte herausgreifen, die mir missfallen haben.

Inhaltsverzeichnis

1 Seite 37: Quantile

- Bei 8 Werten ergibt sich das 0.25-Quantil gemäss "Statistik angewandt" als Wert Nummer 2.5, also als $x_{[2]} + 0.5 \cdot (x_{[3]} - x_{[2]})$.
  Obwohl das Buch sich auf die Statiatikberechnungen in Excel stützt, ist diese Rechnung nicht konform mit der Berechung des Quantils in Excel!
- Excel (wie auch R) liefert Wert Nummer $1 + 0.25 \cdot (8 - 1) = 2.75$, also $x_{[2]} + 0.75 \cdot (x_{[3]} - x_{[2]})$
  Bemerkung: Dieser Ansatz ist wohl dem Bestreben geschuldet, dass man in jedem Fall einen Wert ausweisen will! Dies gelingt, da der Index $1 + q \cdot (8 - 1) \in [1, 8]$ für alle $q \in [0, 1]$
- Mathematisch fundiert wäre folgendes Vorgehen:
  Bezeichne $q(x) := P(X \leq x)$ das Quantil des Wertes x.
  Ziehen wir n Werte und ordnen diese aufsteigend $X_{[1]} \leq X_{[2]} \leq \ldots \leq X_{[n]}$, so ergibt sich die Wahrscheinlichkeitsdichte f_k von $q(X_{[k]})$ zu $f_k(q) = n \cdot \binom{n-1}{k-1} \cdot q^{k-1} \cdot (1-q)^{n-k}$
  1) Kontrolle Regularität:
  $$\int_0^1 n \cdot \binom{n-1}{k-1} \cdot q^{k-1} \cdot (1-q)^{n-k} dq = n \cdot \binom{n-1}{k-1} \cdot B(k, n-k+1) =$$
  $$= \frac{n \cdot (n-1)!}{(k-1)! \cdot (n-k)!} \cdot \frac{\Gamma(k)\Gamma(n-k+1)}{\Gamma(n+1)} = \frac{n!}{(k-1)! \cdot (n-k)!} \cdot \frac{(k-1)!(n-k)!}{n!} = 1$$
  2) Erwartungswert $E[q(X_{[k]})]$:
  $$\int_0^1 q \cdot n \cdot \binom{n-1}{k-1} \cdot q^{k-1} \cdot (1-q)^{n-k} dq = n \cdot \binom{n-1}{k-1} \cdot B(k+1, n-k+1) =$$
  $$= \frac{n \cdot (n-1)!}{(k-1)! \cdot (n-k)!} \cdot \frac{\Gamma(k+1)\Gamma(n-k+1)}{\Gamma(n+2)} = \frac{n!}{(k-1)! \cdot (n-k)!} \cdot \frac{k!(n-k)!}{(n+1)!} = \frac{k}{n+1}$$
  $\Rightarrow 0.25 = \frac{k}{n+1} \Rightarrow$ Wert Nummer $k = 0.25 \cdot (n+1) = 0.25 \cdot 9 = 2.25$, also $x_{[2]} + 0.25 \cdot (x_{[3]} - x_{[2]})$
- Bei den drei Verfahren kommen die 8 Messwerte an den folgenden Quantilen zu liegen:



2 Seite 94: Grundgesamtheit Ω der Grösse $|\Omega| = N < \infty$

- In diesem Zusammenhang werden schwammige und auch unhaltbare Behauptungen aufgestellt!
- So sollte es sein:
- Definiere

  $\mu := \frac{1}{|\Omega|} \cdot \sum_{x \in \Omega} x$

  $\sigma^2 := \frac{1}{|\Omega|} \cdot \sum_{x \in \Omega} (x-\mu)^2.$

  Es wird eine Stichprobe $X_1, X_2, \ldots X_n$ vom Umfang $n \leq N$ aus der Grundgesamtheit gezogen.

  $\overline{X}_n := \frac{1}{n} \cdot \sum_{i=1}^{n} X_i$

  $S_n^2 := \frac{1}{n-1} \cdot \sum_{i=1}^{n} (X_i - \overline{X}_n)^2$

Folgendes lässt sich zeigen:

- Satz

  1) $E(S_n^2) = \frac{N}{N-1} \cdot \sigma^2$

  2) $V(\overline{X}_n) = \frac{N-n}{N-1} \cdot \frac{\sigma^2}{n}$

  3) $V(\overline{X}_n) = \frac{N-n}{N} \cdot \frac{E(S_n^2)}{n}$

- Lemma

  Für $1 \leq i \leq n$, $1 \leq j \leq n$ und $i \neq j$ gilt (jeweils inklusive Beweis):

  ◦ $E(X_i) = \frac{1}{|\Omega|} \cdot \sum_{x \in \Omega} x = \mu$, speziell $E(X_i) = E(X_1)$

  ◦ $V(X_i) = \frac{1}{|\Omega|} \cdot \sum_{x \in \Omega} (x-\mu)^2 = \sigma^2$, speziell $V(X_i) = V(X_1)$

  ◦ $E(X_i^2) = E(X_i - \mu + \mu)^2 = E(X_i - \mu)^2 + 2 \cdot E(X_i - \mu) \cdot \mu + \mu^2 = \sigma^2 + 2 \cdot 0 \cdot \mu + \mu^2 = \sigma^2 + \mu^2$,
  speziell $E(X_i^2) = E(X_1^2)$

  ◦ $E(X_i X_j) = E(E(X_i X_j | X_i)) = E(X_i \cdot \frac{N\mu - X_i}{N-1}) = E(X_i \cdot \frac{N\mu}{N-1}) - E(\frac{X_i^2}{N-1}) = \mu \cdot \frac{N\mu}{N-1} - \frac{\mu^2 + \sigma^2}{N-1} =$
  $= \mu^2 - \frac{\sigma^2}{N-1}$, speziell $E(X_i X_j) = E(X_1 X_2)$

  ◦ $C(X_i, X_j) = E(X_i X_j) - E(X_i)E(X_j) = \mu^2 - \frac{\sigma^2}{N-1} - \mu^2 = -\frac{\sigma^2}{N-1}$, speziell $C(X_i, X_j) = C(X_1, X_2)$

- Beweis von 1):

  $E(S_n^2) = E(\frac{1}{n-1} \cdot \sum_{i=1}^{n} (X_i - \overline{X}_n)^2) = \frac{n}{n-1} \cdot E(\frac{n-1}{n} \cdot X_1 - \frac{1}{n} \cdot \sum_{i=2}^{n} X_i)^2 =$

  $= \frac{n}{n-1} \cdot [(\frac{(n-1)^2}{n^2} + (n-1) \cdot \frac{1}{n^2}) \cdot E(X_1^2) + (-2 \cdot \frac{(n-1)^2}{n^2} + \frac{(n-1) \cdot (n-2)}{n^2}) \cdot E(X_1 X_2)] =$

  $= (\mu^2 + \sigma^2) - (\mu^2 - \frac{\sigma^2}{N-1}) = \frac{N}{N-1} \cdot \sigma^2$ qed.

  Beweis von 2):

  $V(\overline{X}_n) = V(\frac{\sum_{i=1}^{n} X_i}{n}) = \frac{n}{n^2} \cdot V(X_1) + \frac{n(n-1)}{n^2} \cdot C(X_1, X_2) = \frac{1 - \frac{n-1}{N-1}}{n} \cdot \sigma^2 = \frac{N-n}{N-1} \cdot \frac{\sigma^2}{n}$ qed.

  Beweis von 3):

  Aus 2) und 1) folgt: $V(\overline{X}_n) = \frac{N-n}{N-1} \cdot \frac{\sigma^2}{n} = \frac{N-n}{N-1} \cdot \frac{E(S_n^2) \cdot (N-1)}{n \cdot N} = \frac{N-n}{N} \cdot \frac{E(S_n^2)}{n}$ qed.

3 Seite 113: $X \sim N(\mu, \sigma)$

Heisst ganz klar: X ist normalverteilt gemäss $N(\mu, \sigma)$.
Kronthaler spricht von annähernd normalverteilt. Das wäre aber $X \approx N(\mu, \sigma)$

4 Seite 126: Alternativhypothese

$H_0 : \mu = 40$, $H_A = \neg H_0$, also $H_A : \mu \neq 40$
und nicht $H_A : \overline{x} \neq 40$ wie im Buch steht.

Literaturverzeichnis

Kronthaler, Franz (2014): Statistik angewandt.
1. Auflage. Berlin Heidelberg: Springer Verlag