

Zusammenschrift zur Normalverteilung

9. Mai 2015

1 Normalverteilung

1.1 Bestimmte Wahrscheinlichkeit

$$dnorm(x, \mu, \sigma) = \text{Wahrscheinlichkeit für einen bestimmten Wert} \quad (1)$$

1.2 Kumulative Wahrscheinlichkeit

$$pnorm(x, \mu, \sigma) = \text{Wahrscheinlichkeit für höchstens einen Wert} \quad (2)$$

1.3 Inverse kumulative Wahrscheinlichkeit

$$qnorm(p, 0, 1) = x = \text{Wert für eine bestimmte Wahrscheinlichkeit} \quad (3)$$

2 Verschiebung zur Standardnormalverteilung

$$u = \frac{x - \mu}{\sigma} = qnorm(p, 0, 1) \quad (4)$$

3 Normalverteilung und Stichproben

Wenn eine Stichprobe mit n Werten von einer Grundgesamtheit genommen wird, dann verändern sich μ und σ , die eigentlich für die Grundgesamtheit gelten, folgendermaßen:

$$\bar{\mu} = \mu \quad (5)$$

$$\bar{\sigma} = \frac{\sigma}{\sqrt{n}} \quad (6)$$

4 Annäherung der Binomialverteilung zur Normalverteilung

Es gilt:

$$\mu = n * p \quad (7)$$

$$\sigma = \sqrt{n * p * (1 - p)} \quad (8)$$

Es muss erkannt werden, dass die gegebenen Daten binomialverteilt sind! Dies ist gegeben, wenn die Daten immer nur Ja/Nein Ergebnisse sind und bei einer Wiederholung sich die Wahrscheinlichkeit nicht ändert.

5 Suchen nach Variablen anhand einer Binomialverteilung

5.1 Grundlegendes

Mit den im Kapitel section 1 - Normalverteilung beschriebenen Funktionen können nun folgende Werte gesucht werden:

x Der bestimmte Wert von einer Grundgesamtheit

n Die Grundgesamtheit

p Die Wahrscheinlichkeit, wieviel von der Grundgesamtheit y eintritt ($y = n * p$)

p₂ Die Wahrscheinlichkeit einen Wert x von der Grundgesamtheit n zu erhalten

Wichtig ist, dass die folgenden Vorgehensweisen sich nicht auf die Binomialverteilung beschränkt. Es muss nur μ und σ gegeben sein! Für all die nachstehenden Suchen sind diese beide definiert durch:

- $\mu = p * n$
- $\sigma = \sqrt{n * p * (1 - p)}$

5.2 Suche nach x

Gegebene Variablen:

- n
- p
- Phöchstens x Elemente

$$qnorm(p_{\text{höchstens x Elemente}}, 0, 1) = \frac{x - \mu}{\sigma} \text{ solve, } x \rightarrow \quad (9)$$

5.3 Suche nach n

Gegebene Variablen:

- x
- p
- Phöchstens x Elemente

$$qnorm(p_{\text{höchstens x Elemente}}, 0, 1) = \frac{x - \mu}{\sigma} \text{ solve, } n \rightarrow \quad (10)$$

5.4 Suche nach p

Gegebene Variablen:

- x
- n

Wenn es sich um eine Binomialverteilung handelt, dann gilt: $\sigma = \sigma_0$, da eine Wurzel gezogen wurde.

$$pnorm(x, \mu, \sigma) = p_2 \quad (11)$$

6 Zweiseitiger Zufallsstreubereich

6.1 Grundlegendes

Beim zweiseitigen Zufallsstreubereich wird ermittelt, welche Unter- und Obergrenze die Werte in einem bestimmten Bereich symmetrisch verteilt um den Erwartungswert besitzt.

Gegebene Variablen:

μ

σ

p Die Wahrscheinlichkeit (Größe) des Streubereichs symmetrisch um den Erwartungswert μ

6.2 Berechnung von $x_{u,o}$

$$\alpha = 1 - p \quad (12)$$

$$x_u = \mu - qnorm(1 - \frac{\alpha}{2}, 0, 1) * \sigma \quad (13)$$

$$x_o = \mu + qnorm(1 - \frac{\alpha}{2}, 0, 1) * \sigma \quad (14)$$

6.3 Berechnung von $x_{u,o}$ bei Stichproben

Es kommt nun zusätzlich die gegebene Variable der Stichprobenanzahl n hinzu.

$$\alpha = 1 - p \quad (15)$$

$$\bar{\sigma} = \frac{\sigma}{\sqrt{n}} \quad (16)$$

$$x_u = \mu - qnorm(1 - \frac{\alpha}{2}, 0, 1) * \bar{\sigma} \quad (17)$$

$$x_o = \mu + qnorm(1 - \frac{\alpha}{2}, 0, 1) * \bar{\sigma} \quad (18)$$

7 Konfidenzintervall

7.1 Grundlegendes

Das Konfidenzintervall gibt an, in welchen Intervall die Standardabweichung am ehestens bei gegebenen Stichproben liegt. Das bedeutet, dass σ selbst nicht ermittelt werden kann, jedoch der Bereich, in der sich σ bewegt.

Gegeben Variablen müssen sein:

m Ein einzeliger Stichprobenvektor

s Die Standardabweichung s berechnet aus m errechnet mit `Stdev()` bzw. `stdev()`

p Das Konfidenzintervall (meistens 95 %)

7.2 Suche nach dem Konfidenzintervall

Es werden folgende Variablen gesucht:

σ_u Die untere Grenze des Vertrauensbereichs

σ_o Die obere Grenze des Vertrauensbereichs

Berechnung:

$$\alpha = 1 - p \quad (19)$$

$$p_u = 1 - \frac{\alpha}{2} \quad (20)$$

$$p_o = \frac{\alpha}{2} \quad (21)$$

$$\sigma_u = s * \sqrt{\frac{f}{qchisq(p_u, f)}} \quad (22)$$

$$\sigma_o = s * \sqrt{\frac{f}{qchisq(p_o, f)}} \quad (23)$$

Der p %ige Vertrauensbereich für σ liegt dann bei: $[\sigma_u; \sigma_o]$

8 Ermittlung von μ

8.1 Ermittlung basierend auf Stichproben

Es wird, ähnlich wie bei section 7 - Konfidenzintervall, die Erwartungswert μ zu Stichproben in einem Intervall errechnet. Es ist dabei σ gegeben!

Gegebene Variablen:

m Ein einzeliger Stichprobenvektor

n Die Länge von m

σ Die Standardabweichung

α Die Signifikanz in Prozent

8.1.1 Berechnung von μ

$$\bar{x} = \text{mean}(m) \quad (24)$$

$$u = \text{qnorm}\left(1 - \frac{\alpha}{2}, 0, 1\right) \quad (25)$$

$$\mu_u = \bar{x} - u * \frac{\sigma}{\sqrt{n}} \quad (26)$$

$$\mu_o = \bar{x} + u * \frac{\sigma}{\sqrt{n}} \quad (27)$$

8.2 Studentsche Verteilung

8.2.1 Grundlegendes

Es wird der Erwartungswert μ auf Grund einer Stichprobe ermittelt. Hier ist jedoch σ nicht bekannt und muss selbst errechnet werden. Der Rest deckt sich mit Kapitel subsection 8.1 - Ermittlung basierend auf Stichproben.

Gegebene Variablen:

m Ein einzeiliger Stichprobenvektor

n Die Länge von m

s Die Standardabweichung s berechnet aus m errechnet mit *Stdev()* bzw. *stdev()*

α Die Signifikanz in Prozent

8.2.2 Berechnung von μ

$$\bar{x} = \text{mean}(m) \quad (28)$$

$$s = \text{Stdev}(m) \text{ bzw. } \text{stdev}(m) \quad (29)$$

$$f = n - 1 \quad (30)$$

$$\mu_u = \bar{x} - \text{qt}\left(1 - \frac{\alpha}{2}, f\right) * \frac{s}{\sqrt{n}} \quad (31)$$

$$\mu_o = \bar{x} + \text{qt}\left(1 - \frac{\alpha}{2}, f\right) * \frac{s}{\sqrt{n}} \quad (32)$$

9 χ^2 -Verteilung

9.1 Grundlegendes

Die χ^2 -Verteilung wird verwendet um zu Überprüfen:

- Wie wahrscheinlich ist eine bestimmte Standardabweichung s? (Kapitel subsection 9.6.2 - Suche nach p)
- Zu einer Wahrscheinlichkeit p wird welche Standardabweichung s erwartet? (Kapitel subsection 9.6.1 - Suche nach s)
- Eine alternierende Standardabweichung s ist zur Standardabweichung σ gesucht.

Wichtig dabei ist, dass es sich bei der Standardabweichung s um eine Standardabweichung gegeben von bestimmten Werten handelt. Somit wird s entweder durch einen externen Rechner gegeben oder selbst mit den folgenden Funktionen ermittelt:

Stdev(m) m ist ein einzeiliger Vektor mit der Grundgesamtheit

stdev(m) m ist ein einzeiliger Vektor mit Stichproben

9.2 Freiheitsgrade

Die Werte n, die freigewählt werden können. Die Definition der Freiheitsgrade:

$$f = n - 1 \quad (33)$$

9.3 Bestimmte Wahrscheinlichkeitsdichte

$$dchisq(x, f) = \text{Wahrscheinlichkeitsdichte für einen bestimmten Wert} \quad (34)$$

9.4 Kumulative Wahrscheinlichkeitsdichte

$$x_{\text{prüf}} = f * \frac{s^2}{\sigma^2} \quad (35)$$

$$pchisq(x_{\text{prüf}}, f) = \text{Wahrscheinlichkeitsdichte für höchstens einen Wert} \quad (36)$$

9.5 Inverse kumulative Wahrscheinlichkeitsdichte

$$\chi^2_{f,p} = qchisq(p, f) = x_{\text{prüf}} \quad (37)$$

9.6 Suche nach Werten

Es müssen wie im Kapitel subsection 5.1 - Grundlegendes μ und σ gegeben sein. Außerdem müssen die Freiheitsgrade f bekannt sein (\rightarrow also die Anzahl der Werte!).

9.6.1 Suche nach s

Gegebene Variablen:

- p
- Es ist bekannt ob eine maximale (Kapitel subsection 9.4 - Kumulative Wahrscheinlichkeitsdichte) oder eine bestimmte (Kapitel subsection 9.3 - Bestimmte Wahrscheinlichkeitsdichte) Standardabweichung gesucht ist. (Matchad beherrscht nur die inverse kumulative Wahrscheinlichkeitsdichte!)

$$x_{\text{prüf}} = qchisq(p, f) \quad (38)$$

$$x_{\text{prüf}} = f * \frac{s^2}{\sigma^2} \text{ solve, } s \rightarrow \quad (39)$$

9.6.2 Suche nach p

Gegebene Variablen:

- s bzw. m für $Stdev(m)$ oder $stdev(m)$

Es wird somit die Berechnung des Kapitels subsection 9.4 - Kumulative Wahrscheinlichkeitsdichte verwendet.