Studiengang Molekulare Biotechnologie
Biostatistik - MoBi
Sommersemester 2018
Carl Herrmann

– Übungsblatt 4 –

Aufgabe 1 :

Es wird überprüft, ob eine Füllanlage Flaschen tatsächlich mit $v = 1000\,ml$ Saft füllt. Eine Stichprobe von $n = 16$ Flaschen wird untersucht; der Mittelwert der Füllmenge ergibt $\bar{v} = 998.2875\,ml$, mit einer Standardabweichung von $s = 2.5\,ml$.

1. Formulieren Sie die H_0 Hypothese
2. Wird hier ein einseitiger oder zweiseitiger Test durchgeführt ?
3. Berechnen Sie die Teststatistik θ
4. Wir nehmen an, dass die H_0 Verteilung durch die SNV gegeben ist; untersuchen Sie anhand dieser Tabelle (Link), ob die H_0 Hypothese bei einem Signifikanzwert von $\alpha = 0.05$ verworfen werden kann.

Lösung:

1. Wir definieren eine Zufallsvariable V als "Füllmenge"; die H_0 Hypothese lautet dann: "der Erwartungswert von V ist $v = 1000ml$. Die Alternativhypothese H_1 lautet dann "Der Erwartungswert von V ist kleiner oder größer als $v = 1000ml$. Man könnte die H_0 Hypothese auch anders formulieren, z.B. H_0: der Unterschied zwischen dem Mittelwert und dem Sollwert kann durch statistische Fluktuationen erklärt werden". Es wird also ein zweiseitiger Test durchgeführt.

3. Die Teststatistik ist:

$$\theta = \frac{998.2875 - 1000}{\sigma/\sqrt{n}}$$

Hier ist die Größe der Stichprobe $n = 1$ und $\sigma = 2.5$, also ist

$$\theta = -2.74$$

Für einen zweiseitigen Test wäre der kritische Wert $\theta_{95} = -1.96$. Der beobachtete Wert von θ liegt also im Signifikanzbereich (da $|\theta| > |\theta_{95}|$), daher ist der Test signifikant, und die H_0-Hypothese kann verworfen werden.

Aufgabe 2 :

Sie kaufen einen neues VW-Auto, das laut Betriebsanleitung mit einer Tankfüllung (40l) im Schnitt 700 Kilometer fährt. Nach dem Diesel-Skandal sind Sie natürlich VW gegenüber etwas skeptisch geworden...

1. Die 1. Tankfüllung ist nach 696 Kilometern verbraucht. Besteht Anlass zu glauben, dass die Angabe aus der Betriebsanleitung ein Fehler ist? Gehen Sie von einer zu erwartenden Streuung von $\sigma = 3$ (die auch in der Betriebsanleitung angegeben ist!) aus. Formulieren Sie vor Ihren Berechnungen eine Nullhypothese und eine Alternativhypothese. Der Signifikanzbereich liegt bei $\alpha = 0.05$.

2. Daraufhin führen Sie eine kontrollierte Studie mit 5 Durchgängen durch. Jede Tankfüllung fahren Sie bis auf den letzten Tropfen. Berechnen Sie anhand dieser Daten, ob Anlass besteht zu glauben, dass die Angabe aus der Anleitung ein Fehler ist. Gehen Sie dabei davon aus, dass die Streuung nicht bekannt ist und der Standardfehler aus den Daten geschätzt werden muss Formulieren Sie vor Ihren Berechnungen eine Nullhypothese und eine Alternativhypothese. Der Signifikanzbereich liegt auch hier bei $\alpha = 0.05$.

1. Fahrt	2. Fahrt	3. Fahrt	4. Fahrt	5. Fahrt
685	699	681	689	691

Lösung: *Hier definieren wir die Zufallsvariable als S: "Strecke die mit einer Tankfüllung gefahren werden kann". Die H_0 Hypothese lautet hier: "Der Erwartungswert von S ist größer oder gleich 700 km". Die Alternativhypothese H_1 lautet also "der Erwartungswert von S ist kleiner als 700km". Hier macht es mehr Sinn, einen einseitigen Test anzuwenden, da nur Abweichungen nach unten von Interesse sind. Es könnte aber auch ein zweiseitiger Test durchgeführt werden, jedoch wäre es dann schwieriger, ein signifikantes Ergebnis zu bekommen.*

Wir berechnen die Teststatistik:

$$t = \frac{696 - 700}{\sigma/\sqrt{n}}$$

wo $n = 1$ ist und die Standardabweichung mit $\sigma = 3$ angegeben wird (mit einer Stichprobe $n = 1$ kann mit bestem Willen keine Standardabweichung geschätzt werden...) Also ist $t = -1.33$. Für einen einseitigen Test wäre der kritische Wert einer SNV $\theta_{95} = -1.64$, also ist der Test hier nicht signifikant, da θ nicht im Signifikanzbereich liegt! Wir können einen Betrug also nicht eindeutig feststellen...

Nach den $n = 5$ Testfahrten können wir Mittelwert und Stichprobenvarianz berechnen (da wir die Stichprobenvarianz als Schätzer für die Varianz benutzen benutzen wir den Faktor $\frac{1}{N-1}$).

$$m = 689; s = 6.78233$$

Daraus leiten wir ab

$$\theta = \frac{689 - 700}{s/\sqrt{5}} = -3.626593$$

Hier ist der Test eindeutig signifikant, auch wenn wir anstatt der SNV als H_0 Verteilung die t-Verteilung mit 4 Freiheitsgraden annehmen. In diesem Fall ist der kritische Wert $t_{95} = -2.12$, und der beobachtete Wert der Teststatistik liegt weiterhin im Signifikanzbereich. Also können wir H_0 verwerfen, und es liegt tatsächlich eine Kundentäuschung vor.