

Bachelorprüfung: Statistik Musterlösung

1. a) $P(X > 104) = 0.345$
 b) $P(\bar{X} > 104) = 0.023$
 c) $P(\bar{X} \geq 104) = 0.023$, d.h., die Nullhypothese wird auf dem 5% Niveau verworfen.
 d) Verwerfungsbereich: $[103.3; \infty)$
 e) Fehler zweiter Art: $P(\bar{X}_{\mu=105} \leq 103.3) = 0.198$.

2. 1) Wie gross ist σ ? b) 36.6

2) Wie lautet die korrekte Null-, respektive Alternativ-Hypothese?

- a) $H_0 : \mu_0 = 100$
 $H_1 : \mu_1 > 100$

3) Welche Verteilung hat \bar{X}_n unter der Nullhypothese?

- b) $\mathcal{N}(\mu_0, \sigma^2/n)$

4) Nehmen Sie an $\bar{X}_{25} = 115$. Testen Sie H_0 gegen H_1 , indem Sie einen z-Test auf dem 5% Niveau durchführen.

- b) Der p -Wert beträgt 0.016, daher wird die Nullhypothese auf dem 5%, nicht aber auf dem 1% Niveau verworfen.

$$\text{Aus } \frac{\bar{X}_{25} - 100}{\sigma/\sqrt{25}} \sim \mathcal{N}(0, 1) \quad \text{folgt:}$$

$$p - \text{Wert} = 1 - \Phi\left(\frac{115 - 100}{35/5}\right) = 1 - \Phi(15/7) = 0.016$$

5) Geben Sie explizit den Annahmehereich des Testes auf dem 5% Niveau an.

- b) $[-\infty, 111.51]$

Der Annahmehereich ist einseitig, daher gilt:

$$[-\infty, \Phi^{-1}(0.95)\frac{\sigma}{\sqrt{n}}] = [-\infty, \Phi^{-1}(0.95)\frac{35}{5}] = [-\infty, 111.51].$$

6) Wie gross ist die Macht des Testes in Teilaufgabe 4, falls die IQ's von Gymnasiasten $\mathcal{N}(\mu, 35^2)$ mit $\mu = 120$ verteilt ist?

b) 0.89

Es gilt $\bar{X}_{25} \sim \mathcal{N}(120, 35^2/25)$. Die Macht des Testes ist der folgende Ausdruck:

$$\begin{aligned} P [111.51 \leq \bar{X}_{25} \leq \infty] &= P \left[\frac{111.51 - 120}{35/5} \leq \frac{\bar{X}_{25} - 120}{35/5} \leq \infty \right] \\ &= 1 - \Phi\left(\frac{111.51 - 120}{35/5}\right) = 0.89 \end{aligned}$$

7) Wie verändert sich die Macht des Testes, falls das wahre μ grösser als 120 ist?

a) Sie wird grösser.

8) Der Fehler 2. Art wird kleiner, falls das Niveau α verkleinert wird.

b) Stimmt nie.

9) Welche der folgenden Aussagen ist richtig (nur eine stimmt)?

d) Liegt μ_0 im Vertrauensbereich, so ist der p -Wert des zugehörigen Tests nicht signifikant.

3. a) Binomialverteilt: $X \sim \mathcal{B}(10, 0.1)$.

b) Kein erkrankter Fuchs wird gefangen: $P(X = 0) = 0.9^{10} \approx 0.348$. Mindestens zwei erkrankte Fuchse werden gefangen:

$$P(X \geq 2) = 1 - P(X = 0) - P(X = 1).$$

Es gilt $P(X = 1) = 10 \cdot 0.9^9 \cdot 0.1^1 \approx 0.387$. Und somit

$$P(X \geq 2) \approx 0.264.$$

c) Approximation durch Normalverteilung:

$$X \sim \mathcal{N}(\mu, \sigma^2).$$

mit $\mu = 200 \cdot 0.1 = 20$ und $\sigma^2 = 200 \cdot 0.9 \cdot 0.1 = 18$. Mit der Normalapproximation gilt $P(X \leq 26) < 0.95$ und $P(X \leq 27) > 0.95$. Also $q = 27$.

d) Die Anzahl 30 erkrankter Füchse ist grösser als $q = 27$. Die Nullhypothese dass $p \leq 0.1$ kann somit verworfen werden.

4. 1) a)

2) a)

3) f)

4) c)

5) a)

- 6) b)
- 7) d)
- 8) b)
- 9) c)
- 10) a)
- 11) f)

- 5.
- 1) d)
 - 2) b)
 - 3) a)
 - 4) d)
 - 5) b)

- 6) b) Zur Erklärung: Es werden 25 von 100 Personen als Raucher beobachtet. Man interessiert sich für einen einseitigen Test und möchte wissen, ob der Anteil der Raucher 0.3 oder mehr ist. Das, was interessiert steckt man in die Alternative $H_A : p \geq 0.3$. Die Nullhypothese ist dann $H_0 : p \leq 0.3$.

Gesucht ist nun ein Vertrauensintervall für den Anteil p der Raucher in der Bevölkerung. Um ein Vertrauensintervall zu erzeugen, geht man alle möglichen Parameterwerte p_0 durch und fragt sich bei jedem Parameterwert: "Kann die Nullhypothese $H_0 : p \leq p_0$ beibehalten werden (Beachte: Die Form $p \leq p_0$ ist genau wie im Test, der im letzten Abschnitt erklärt wurde), d.h., passt der Wert p_0 zu den beobachteten Daten?"

Gehen wir ein paar Werte durch:

- $p_0 = 1$: 25 Raucher unter 100 passt zu der Behauptung in $H_0 : p \leq 1$ ("Es rauchen 100% der Bevölkerung oder weniger"). Also liegt $p_0 = 1$ im Vertrauensintervall. Das gleiche Argument gilt auch für andere grosse Werte von p_0 .
- $p_0 = 0$: 25 Raucher unter 100 passt nicht zu der Behauptung in $H_0 : p \leq 0$ ("Es raucht niemand in der Bevölkerung"). Also liegt $p_0 = 0$ nicht im Vertrauensintervall. Das gleiche Argument gilt auch für andere kleine Werte von p_0 .
- Allgemein passt die Beobachtung also zu grossen Werten von p_0 , nicht aber zu kleinen. Irgendwo dazwischen muss also die Grenze liegen. Liegt sie nun über oder unter 0.25? Angenommen, $p_0 = 0.26$ und $H_0 : p \leq p_0$, dann ist es durchaus plausibel 25 von 100 "Treffer" zu beobachten. H_0 wird also beibehalten und $p_0 = 0.26$ liegt im Vertrauensintervall. Wenn p_0 etwas kleiner wird, z.B. $p_0 = 0.23$, dann hat man zwar etwas mehr "Treffer" beobachtet als erwartet (man hätte ja 23 Treffer erwartet und 25 beobachtet), allerdings könnte dieses Ergebnis noch ohne Weiteres per Zufall entstanden sein und $p_0 = 0.23$ wirklich richtig sein. Irgendwo zwischen 0 und einer Zahl kleiner als 0.25 muss also die Untergrenze des Vertrauensintervalls liegen.

Damit ist die einzig plausible Lösung in der Auswahl die Antwort (b): $[0.18, 1]$ oder $[18\%, 100\%]$.