Wahrscheinlichkeitsrechnung und Statistik für Studierende der Informatik

Vorlesung 14

Konfidenzintervalle

Bisher haben wir versucht einen unbekannten Parameter θ der Verteilung X durch eine konkrete Stichprobe punktgenau zu schätzen.

Mehr Erfolg haben wir natürlich, wenn wir „nur“ schätzen in welchem Intervall θ liegt.

Idee: Wir möchten anhand einer konkreten Stichprobe $(x_1, \ldots, x_2)$ von X ein Intervall $[\hat{\theta}_u, \hat{\theta}_o]$ bestimmen, so dass θ mit einer von uns vorgegebenen Wahrscheinlichkeit $1 - \alpha$ in diesem Intervall liegt.

ABER: θ ist fest, nur unbekannt!

GENAUER: Wir suchen zwei Schätzfunktionen $\hat{\theta}_u(X_1, \ldots, X_n)$ und $\hat{\theta}_o(X_1, \ldots, X_n)$, so dass $P(\hat{\theta}_u \leq \theta \leq \hat{\theta}_o) = 1 - \alpha$, für vorgegebenes $\alpha \in (0,1)$.

NOCH GENAUER: Wir suchen zwei Schätzfunktionen $\hat{\theta}_u(X_1, \ldots, X_n)$ und $\hat{\theta}_o(X_1, \ldots, X_n)$,

so dass $P_\theta(\hat{\theta}_u \leq \theta \leq \hat{\theta}_o) = 1 - \alpha$, für vorgegebenes $\alpha \in (0,1)$ und alle möglichen θ!

Konfidenzintervalle

Definition 5.19. Ist θ ein zu schätzender Parameter und sind $\hat{\theta}_u$ und $\hat{\theta}_o$ Schätzfunktionen mit der Eigenschaft

$$P_\theta(\hat{\theta}_u \leq \theta \leq \hat{\theta}_o) = 1 - \alpha \text{ (oder auch } \geq 1 - \alpha)$$

- so liefert jede konkrete Stichprobe $x = (x_1, \ldots, x_n)$ ein *Konfidenzintervall* $[\hat{\theta}_u(x), \hat{\theta}_o(x)]$.
- Die Zahl $1 - \alpha$ heißt *Konfidenzniveau*. Die Zahl α heißt *Irrtumswahrscheinlichkeit.*

Häufig ist man in der Situation, die *unteren Konfidenz- oder Vertrauensgrenze* und die *obere Konfidenz- oder Vertrauensgrenze* für einen Parameter θ einzeln zu bestimmmen:

- $\hat{\theta}_u$ heißt *untere Konfidenzgrenze* für den Parameter θ zum Konfidenzniveau $1 - \beta$, wenn
$$P_\theta(\hat{\theta}_u \leq \theta) \geq 1 - \beta.$$
- $\hat{\theta}_o$ heißt *obere Konfidenzgrenze* für den Parameter θ zum Konfidenzniveau $1 - \beta$, wenn
$$P_\theta(\theta \leq \hat{\theta}_o) \geq 1 - \beta.$$

Konfidenzintervalle

Häufig ist man in der Situation, die *unteren Konfidenz- oder Vertrauensgrenze* und die *obere Konfidenz- oder Vertrauensgrenze* für einen Parameter θ einzeln zu bestimmmen:

- $\hat{\theta}_u$ heißt *untere Konfidenzgrenze* für den Parameter θ zum Konfidenzniveau $1-\beta$, wenn
$$P_\theta(\hat{\theta}_u \leq \theta) \geq 1-\beta.$$
- $\hat{\theta}_o$ heißt *obere Konfidenzgrenze* für den Parameter θ zum Konfidenzniveau $1-\beta$, wenn
$$P_\theta(\theta \leq \hat{\theta}_o) \geq 1-\beta.$$

Bemerkung: Aus $P_\theta(\hat{\theta}_u \leq \theta) \geq 1-\frac{\alpha}{2}$ und $P_\theta(\theta \leq \hat{\theta}_o) \geq 1-\frac{\alpha}{2}$ folgt

$$P_\theta(\hat{\theta}_u \leq \theta \leq \hat{\theta}_o) \geq 1-\alpha.$$

Die Notation P_θ bedeutet, dass wir die Wahrscheinlichkeit unter der Annahme, dass der Parameter θ ist, berechnen.

Normal-Verteilung

$$\Phi(x) = \int_{-\infty}^{x} \frac{1}{\sqrt{2\pi}} \cdot \exp\left(-\frac{1}{2}t^2\right) \mathrm{d}t$$

Bemerkung 3.13. Häufig ist man an der Wahrscheinlichkeit interessiert, dass die $N(\mu, \sigma^2)$-verteilte Zufallsvariable Werte in einem zu μ symmetrischen Intervall $[\mu - k\sigma, \mu + k\sigma]$ mit $k \in \mathbb{N}$ annimmt. Es ist üblich, die Abweichung von μ in Einheiten von σ anzugeben. Deshalb spricht man vom *$k\sigma$-Intervall*. Wir erhalten aus den Bemerkungen 3.10:

$$P\big(\mu - k\sigma \le X \le \mu + k\sigma\big) = \phi(k) - \phi(-k) = 2\phi(k) - 1\ .$$

Speziell für $k = 1, 2, 3$ ergeben sich folgende Werte

$$P(\mu - 1 \cdot \sigma \le X \le \mu + 1 \cdot \sigma) = 2\phi(1) - 1 \approx 0.6826\ ,$$

$$P(\mu - 2 \cdot \sigma \le X \le \mu + 2 \cdot \sigma) = 2\phi(2) - 1 \approx 0.9544\ ,$$

$$P(\mu - 3 \cdot \sigma \le X \le \mu + 3 \cdot \sigma) = 2\phi(3) - 1 \approx 0.9974\ .$$

Also liegen ca. 68 % der beobachteten Werte bei einer $N(\mu, \sigma^2)$-verteilten Zufallsvariablen zwischen $\mu - \sigma$ und $\mu + \sigma$, ca. 95 % liegen zwischen $\mu - 2\sigma$ und $\mu + 2\sigma$ und ca. 99.7 % liegen zwischen $\mu - 3\sigma$ und $\mu + 3\sigma$.

Grenzwertsätze

$$\Phi(x) = \int_{-\infty}^{x} \frac{1}{\sqrt{2\pi}} \cdot \exp\left(-\frac{1}{2}t^2\right) \mathrm{d}t$$

Folgerung 4.7. Betrachten wir in Satz 4.6 das symmetrische Intervall $[-k, k]$, so erhalten wir wegen $E(S_n) = n \cdot E(X_1) = n \cdot \mu$ und $Var(S_n) = n \cdot Var(X_1) = n\sigma^2$ die Beziehung

$$\lim_{n\to\infty} P\left(-k \leq \frac{S_n - E(S_n)}{\sqrt{Var(S_n)}} \leq k\right) = \phi(k) - \phi(-k) = 2 \cdot \phi(k) - 1\ ,$$

d.h.

$$\lim_{n\to\infty} P\left(E(S_n) - k\sqrt{Var(S_n)} \leq S_n \leq E(S_n) + k\sqrt{Var(S_n)}\right) = 2 \cdot \phi(k) - 1\ .$$

S_n

Mit Bemerkung 3.13 erhalten wir also, dass die Summe von n unabhängigen und identisch verteilten Zufallsvariablen (als Faustregel für große n) mit einer ungefähren Wahrscheinlichkeit von

- 0.6826 in den Grenzen $E(S_n) \pm 1 \cdot \sqrt{Var(S_n)}$
- 0.9544 in den Grenzen $E(S_n) \pm 2 \cdot \sqrt{Var(S_n)}$
- 0.9974 in den Grenzen $E(S_n) \pm 3 \cdot \sqrt{Var(S_n)}$

liegt.

Konfidenzintervalle

Sei Y eine Zufallsvariable, die die Treffer bei einer n-fachen Wiederholung eines Bernoulli-Experimentes zählt. Gesucht ist ein Konfidenzintervall für die Wahrscheinlichkeit p des Bernoulli-Experimentes.

Konkrete Stichprobe: $Y = k$
Geforderte Genauigkeit: $1 - \alpha$

Bestimme den Wert $c = \Phi^{-1}(1 - \alpha/2)$ aus einer Tabelle

$1-\beta$	0.900	0.950	0.975	0.990	0.995
$c = \phi^{-1}(1-\beta)$	1.282	1.645	1.960	2.326	2.576

$\left(\text{mit } \beta = \frac{\alpha}{2}\right)$

Für großes n ergibt sich

Wald-Intervall

$$\hat{p}_u(k) \approx \frac{k}{n} - c \cdot \sqrt{\frac{k(n-k)}{n^3}}$$

$$\hat{p}_o(k) \approx \frac{k}{n} + c \cdot \sqrt{\frac{k(n-k)}{n^3}}$$

Konfidenzintervalle

Sei Y eine Zufallsvariable, die die Treffer bei einer n-fachen Wiederholung eines Bernoulli-Experimentes zählt. Gesucht ist ein Konfidenzintervall für die Wahrscheinlichkeit p des Bernoulli-Experimentes.

Konkrete Stichprobe: $Y = k$

Geforderte Genauigkeit: $1 - \alpha$

$$P_p(Y \leq k) \approx \Phi\left(\frac{k - np + 0{,}5}{\sqrt{np(1-p)}}\right) \approx 1 - \frac{\alpha}{2}$$

Bestimme den Wert $c = \Phi^{-1}(1 - \alpha/2)$ aus einer Tabelle

$1-\beta$	0.900	0.950	0.975	0.990	0.995
$c = \phi^{-1}(1-\beta)$	1.282	1.645	1.960	2.326	2.576

$\left(\text{mit } \beta = \frac{\alpha}{2}\right)$

Dann gilt:

$$\hat{p}_o(k) \approx \frac{k + \frac{1}{2} + \frac{c^2}{2} + c \cdot \sqrt{k + \frac{1}{2} - \frac{1}{n}(k + \frac{1}{2})^2 + \frac{c^2}{4}}}{n + c^2}$$

Wilson-Intervall mit Stetigkeitskorrektur

$$\hat{p}_u(k) \approx \frac{k - \frac{1}{2} + \frac{c^2}{2} - c \cdot \sqrt{k - \frac{1}{2} - \frac{1}{n}(k - \frac{1}{2})^2 + \frac{c^2}{4}}}{n + c^2}\,.$$

Hypothesentest

Tea tasting lady

Manche Behauptungen lassen sich nicht definitiv beweisen oder widerlegen. Man kann aber Versuchen, die Plausibilität der Behauptung statistisch zu überprüfen.

Ein klassisches Beispiel:

Eine britische Lady behauptet sie könne am Geschmack erkennen, ob erst Milch und dann Tee in die Tasse gegossen wurde, oder umgekehrt.

Da das sehr schwierig ist, könne sie allerdings nicht immer richtig liegen.

Wir wollen diese Behauptung überprüfen.

Hypothesentest

Tea tasting lady

Eine britische Lady behauptet sie könne am Geschmack erkennen, ob erst Milch und dann Tee in die Tasse gegossen wurde, oder umgekehrt.

Da das sehr schwierig ist, könne sie allerdings nicht immer richtig liegen.

Wir wollen diese Behauptung überprüfen.

Versuchsaufbau:

1. **Vermutung (Hypothese) aufstellen**
2. **Stichprobe sammeln**
3. **Auswerten: Wie wahrscheinlich ist es, dass diese Stichprobe unter unserer Vermutung auftaucht?**

Konkret:

1. H_0**: Die Lady rät**
2. **Wir servieren ihr 20 Tassen Tee mit Milch. Wir wissen, ob erst Milch oder erst Tee eingeschüttet wurde. Sie muss sich bei jeder Tasse entscheiden und hat** k **Treffer.**
3. **Die Wahrscheinlichkeit für** k **Treffer ist**
$$P(X = k) = \binom{20}{k} \left(\frac{1}{2}\right)^k \left(\frac{1}{2}\right)^{20-k}$$

Hypothesentest

Tea tasting lady

Versuchsaufbau:

1. **Vermutung (Hypothese) aufstellen**
2. **Stichprobe sammeln**
3. **Auswerten: Wie wahrscheinlich ist es, dass diese Stichprobe unter unserer Vermutung auftaucht?**

Konkret:

1. H_0: **Die Lady rät**
2. **Wir servieren ihr 20 Tassen Tee mit Milch. Wir wissen, ob erst Milch oder erst Tee eingeschüttet wurde. Sie muss sich bei jeder Tasse entscheiden und hat k Treffer.**
3. **Die Wahrscheinlichkeit für k Treffer ist**
$$P(X = k) = \binom{20}{k} \left(\frac{1}{2}\right)^k \left(\frac{1}{2}\right)^{20-k}$$

Intuitiv:

- **Hat die Lady $k \leq 10$ Treffer, sollten wir annehmen, dass Sie nur rät.**
- **Hat die Lady $k = 20$ Treffer, sollten wir annehmen, dass Sie die Reihenfolge tatsächlich erschmecken kann.**

Vorsicht:

- **Vielleicht kann Sie es, hatte aber einfach einen schlechten Tag!**
- **Vielleicht hat Sie nur durch raten alle Tassen richtig benannt!**

Hypothesentest

Tea tasting lady

Versuchsaufbau:

1. **Vermutung (Hypothese) aufstellen**
2. **Stichprobe sammeln**
3. **Auswerten: Wie wahrscheinlich ist es, dass diese Stichprobe unter unserer Vermutung auftaucht?**

Konkret:

1. H_0**: Die Lady rät**
2. **Wir servieren ihr 20 Tassen Tee mit Milch. Wir wissen, ob erst Milch oder erst Tee eingeschüttet wurde. Sie muss sich bei jeder Tasse entscheiden und hat** k **Treffer.**
3. **Die Wahrscheinlichkeit für** k **Treffer ist**
$$P(X = k) = \binom{20}{k} \left(\frac{1}{2}\right)^{k} \left(\frac{1}{2}\right)^{20-k}$$

Was ist eigentlich unsere Unbekannte?

Antwort: Die Wahrscheinlichkeit p**, mit der die Lady die Einschüttreihenfolge einer Tasse Tee mit Milch richtig benennt.**

Es gilt $p \in \left[\frac{1}{2}, 1\right]$**. Denn, wenn Sie keinen unterschied Schmecken kann, ist** $p = \frac{1}{2}$ **(sie rät). Anderenfalls wird** p **natürlich größer.**

Hypothesentest

Tea tasting lady

Versuchsaufbau:

1. **Vermutung (Hypothese) aufstellen**
2. **Stichprobe sammeln**
3. **Auswerten: Wie wahrscheinlich ist es, dass diese Stichprobe unter unserer Vermutung auftaucht?**

Unbekannte: $p \in \left[\frac{1}{2}, 1\right]$

Unsere Hypothese $H_0: \; p = \frac{1}{2}$

Alternativhypothese: $H_1: \; p \in \left(\frac{1}{2}, 1\right]$

Mögliche Stichproben: $k \in \{0, \ldots, 20\}$

Konkret:

1. H_0**: Die Lady rät**
2. **Wir servieren ihr 20 Tassen Tee mit Milch. Wir wissen, ob erst Milch oder erst Tee eingeschüttet wurde. Sie muss sich bei jeder Tasse entscheiden und hat** k **Treffer.**
3. **Die Wahrscheinlichkeit für** k **Treffer ist**
$$P(X = k) = \binom{20}{k} \left(\frac{1}{2}\right)^k \left(\frac{1}{2}\right)^{20-k}$$

Gesucht: $k' \in \{0, \ldots, 20\}$**, mit**

$k \leq k'$**, dann sollte** H_0 **gelten**

$k > k'$**, dann sollte** H_1 **gelten**

Hypothesentest

Definition 5.26. Wir zerlegen den Parameterbereich Θ in zwei nichtleere disjunkte Teile

$$\Theta = \Theta_0 \cup \Theta_1 \; .$$

Ein *statistischer Test* oder *Hypothesentest* ist eine Entscheidungsregel, die innerhalb des vorgegebenen Modell-Rahmens für jede mögliche Stichprobe x festlegt, ob man sich für die

$$\textit{Nullhypothese} \;\; H_0 : \quad \text{es gilt } \vartheta \in \Theta_0$$

oder für die

$$\textit{Gegenhypothese (Alternative)} \;\; H_1 : \quad \text{es gilt } \vartheta \in \Theta_1$$

entscheidet. Die Zerlegung von Θ impliziert eine Zerlegung

$$\mathcal{K}_0 \cup \mathcal{K}_1$$

des Stichprobenraums und damit folgende Entscheidungsregel:

$$\text{"Ist } x \in \mathcal{K}_0, \text{ so entscheide für } H_0.\text{"}$$

$$\text{"Ist } x \in \mathcal{K}_1, \text{ so entscheide für } H_1.\text{"}$$

Man sagt auch: "Zu testen ist die Hypothese H_0 gegen die Alternative H_1". $\mathcal{K}_0$ heißt *Annahmebereich* des Tests und $\mathcal{K}_1$ heißt *Verwerfungsbereich* oder auch *kritischer Bereich*. Die Hypothese H_0 wird häufig auch *Nullhypothese* genannt.

Hypothesentest

Tea tasting lady

1. H_0: **Die Lady rät**
2. **Wir servieren ihr 20 Tassen Tee mit Milch. Wir wissen, ob erst Milch oder erst Tee eingeschüttet wurde. Sie muss sich bei jeder Tasse entscheiden und hat k Treffer.**
3. **Die Wahrscheinlichkeit für k Treffer ist** $P(X=k)=\binom{20}{k}\left(\frac{1}{2}\right)^k\left(\frac{1}{2}\right)^{20-k}$

Unbekannte: $p \in \left[\frac{1}{2},1\right]$

Unsere Hypothese $H_0: \ p=\frac{1}{2}$

Eine sichere Entscheidung ist kaum möglich. Daher ist das Ziel: Kann die Hypothese H_0 aus gutem Grund abgelehnt werden oder nicht?

Wir wollen H_0 nur dann ablehnen, wenn die Wahrscheinlichkeit für die Korrektheit von H_0 (anhand unserer Stichprobe k) kleiner oder gleich $\alpha = 0{,}05$ ist.

Dieses α heißt Signifikanzniveau und kann beliebig gewählt werden.

Hypothesentest

Tea tasting lady

1. H_0: **Die Lady rät**
2. **Wir servieren ihr 20 Tassen Tee mit Milch. Wir wissen, ob erst Milch oder erst Tee eingeschüttet wurde. Sie muss sich bei jeder Tasse entscheiden und hat k Treffer.**
3. **Die Wahrscheinlichkeit für k Treffer ist** $P(X = k) = \binom{20}{k}\left(\frac{1}{2}\right)^k\left(\frac{1}{2}\right)^{20-k}$

$p \in \left[\frac{1}{2},1\right]$ $\qquad H_0: \; p = \frac{1}{2}$ $\qquad \alpha = 0{,}05$

Unsere Lady benennt 14 Tassen korrekt. Es gibt folgende Möglichkeiten:

- **Sie rät und wir behaupten auch, dass sie rät.**
- **Sie rät nicht und wir behaupten, dass sie nicht rät.**
- **Sie rät und wir behaupten, dass sie nicht rät**
- **Sie rät nicht und wir behaupten, dass sie rät.**

Fehler 1. Art: H_0 **ist korrekt, aber wir entscheiden uns für** H_1

Fehler 2. Art: H_0 **ist nicht korrekt, aber wir entscheiden uns für** H_0

Hypothesentest

Tea tasting lady

1. H_0: **Die Lady rät**
2. **Wir servieren ihr 20 Tassen Tee mit Milch. Wir wissen, ob erst Milch oder erst Tee eingeschüttet wurde. Sie muss sich bei jeder Tasse entscheiden und hat k Treffer.**
3. **Die Wahrscheinlichkeit für k Treffer ist** $P(X=k) = \binom{20}{k}\left(\frac{1}{2}\right)^k\left(\frac{1}{2}\right)^{20-k}$

$p \in \left[\frac{1}{2}, 1\right]$ $\qquad H_0: \ p = \frac{1}{2}$ $\qquad \alpha = 0{,}05$

Fehler 1. Art: H_0 **ist korrekt, aber wir entscheiden uns für** H_1

Fehler 2. Art: H_0 **ist nicht korrekt, aber wir entscheiden uns für** H_0

Die Lady entscheidet sich bei $k = 14$ Tassen korrekt. Wofür entscheiden wir uns?

Ist H_0 korrekt, so ist $P(X = 14) = 0{,}036...$ **und** $P(X \geq 14) = 0{,}057...$

Die Wahrscheinlichkeit durch Raten, mindestens 14 Treffer zu erzielen ist also $> \alpha = 0{,}05$. Bei 14 Treffern können wir daher nicht mit unserem vorgegebenen Signifikanzniveau α, die Hypothese H_0 verwerfen.

Hypothes

Definition 5.27. Wir übernehmen die Bezeichnungen aus Definition 5.26. Gilt (in Wirklichkeit) $\vartheta \in \Theta_0$ und man entscheidet sich für die Gegenhypothese, so spricht man von einem *Fehler erster Art*. Gilt dagegen $\vartheta \in \Theta_1$ und man entscheidet sich für die Nullhypothese, so spricht man von einem *Fehler zweiter Art*. Man kann dies folgendermaßen zusammenfassen:

		Wirklichkeit	
		$H_0 : \vartheta \in \Theta_0$	$H_1 : \vartheta \in \Theta_1$
Entschei-dung	für H_0	richtige Entscheidung	Fehler 2. Art
	für H_1	Fehler 1. Art	richtige Entscheidung

Um die Wahrscheinlichkeit für eine falsche Entscheidung möglichst klein zu halten, ist eine sog. *Gütefunktion* g mit kleinen Werten auf Θ_0 und großen Werten auf Θ_1 wünschenswert. Wir definieren $g : \Theta \to [0,1]$ durch

$$g(\vartheta) := p_\vartheta(X \in \mathcal{K}_1)\ , \quad \vartheta \in \Theta.$$

g ordnet jedem ϑ die sog. *Verwerfungswahrscheinlichkeit der Hypothese* H_0 *unter* p_ϑ zu. Man gibt eine obere Schranke $\alpha \in]0,1[$ für die Wahrscheinlichkeit des Fehlers erster Art vor und legt $\mathcal{K}_1$ so fest, dass

$$g(\vartheta) \leq \alpha \quad \text{für alle } \vartheta \in \Theta_0$$

gilt. Ein solcher Test heißt *(Signifikanz-)Test zum (Signifikanz-)Niveau* α. Dabei sind für α Werte aus dem Intervall $[0.01, 0.1]$ üblich.

Hypothesentest

„Wir sind jetzt mit dabei,
In jedem siebten Ei!“

Beispiel Ü-Ei: Ist wirklich in jedem siebten Ei eine lustige Figur?

Sei p die Wahrscheinlichkeit für eine Figur im Ei. Dann ist $H_0: \; p \geq \frac{1}{7}$

Gegenhypothese: $H_1: \; p < \frac{1}{7}$

Wir testen 1000 Eier. Wie viele Figuren dürfen höchstens in diesen Eiern sein, damit wir H_0 mit einem Signifikanzniveau von $\alpha = 0.025$ verwerfen können?

Sei X die Zufallsvariable, die die Anzahl von Figuren misst.
Gesucht ist ein $c \in \{0,\ldots,1000\}$, mit $P(X \leq c$ und H_0 stimmt$) \leq \alpha$.