SS 2013

Diskrete Wahrscheinlichkeitstheorie

Javier Esparza

Fakultät für Informatik
TU München

http://www7.in.tum.de/um/courses/dwt/ss13

Sommersemester 2013

Teil V

Induktive Statistik

Induktive Statistik

Das Ziel der induktiven Statistik besteht darin,

- aus gemessenen Zufallsgrößen (z.B. Häufigkeiten)
- auf die zugrunde liegenden Gesetzmäßigkeiten (z.B. W'keiten)

zu schließen.

20. Schätzvariablen

Schätzvariablen I: Einführendes Beispiel

Sei X die Anzahl von Lesezugriffen auf eine Festplatte bis zum ersten Lesefehler.

Wir nehmen $X \sim Geo(p)$ an mit unbekannter Erfolgsw'keit p.
Die W'keit p soll empirisch geschätzt werden.

Wegen $p = \frac{1}{\mathbb{E}[X]}$ schätzen wir $\mathbb{E}[X]$.

Wir führen n Messungen der Anzahl von Zugriffen bis zum ersten Lesefehler. Sei X_i die Zufallsvariable für das Ergebnis der i-ten Messung.

$X_1, \ldots, X_n$ sind unabhängig und besitzen jeweils dieselbe Verteilung wie X.

Die Variablen $X_1, \ldots, X_n$ nennt man Stichprobenvariablen.

Schätzvariablen II: Einführendes Beispiel

Sei $\overline{X}$ das arithmetische Mittel der X_i, d.h. $\overline{X} := \frac{1}{n} \sum_{i=1}^{n} X_i$.

Wir schätzen $\mathbb{E}[X]$ durch $\overline{X}$.

Wir nennen $\overline{X}$ einen Schätzer für den Erwartungswert $\mathbb{E}[X]$.

Schätzvariablen III: Schätzer

Definition 25

Sei X eine Zufallsvariable mit Dichte $f_X(x) = f(x;\theta)$.

Eine Schätzvariable oder kurz Schätzer für θ ist eine Zufallsvariable U der Gestalt $U = f(x_1, \ldots, X_n)$, wobei die $X_1, \ldots, X_n$ unabhängige Stichprobenvariablen sind mit derselben Verteilung wie X.

Frage: Welche Eigenschaften soll eine Schätzvariable erfüllen?

Schätzvariablen III: Erwartungstreue

Definition 26

Ein Schätzer U eines Parameters θ heißt erwartungstreu, wenn $\mathbb{E}[U] = \theta$ gilt.

Die Größe $\mathbb{E}[U - \theta]$ nennt man Bias der Schätzvariablen U.

Satz 27

Sei X eine beliebige Zufallsvariable. $\overline{X}$ is ein erwartungstreuer Schätzer für $\mathbb{E}[X]$.

Beweis:

$$\mathbb{E}[\overline{X}] = \frac{1}{n} \sum_{i=1}^{n} \mathbb{E}[X_i] = \frac{1}{n} \sum_{i=1}^{n} \mathbb{E}[X] = \mathbb{E}[X].$$

$\square$

Schätzvariablen IV: Mittlere quadratische Abweichung

Definition 28

Sei U ein Schätzer eines Parameters θ. Die mittlere quadratische Abweichung von U ist

$$MSE := \mathbb{E}[(U - \theta)^2]\ .$$

Ist U erwartungstreu, dann $MSE = \mathbb{E}[(U - \mathbb{E}[U])^2] = \mathrm{Var}[U]$.

Eine Schätzvariable heißt konsistent im quadratischen Mittel, wenn $MSE \to 0$ für $n \to \infty$ gilt (mit n Umfang der Stichprobe).

Schätzvariablen V: Mittlere quadratische Abweichung

Satz 29

Sei X eine Zufallsvariable mit endlicher Varianz. Der Schätzer $\overline{X}$ für $\mathbb{E}[X]$ ist konsistent im quadratischen Mittel.

Beweis:

Wegen der Unabhängigkeit von $X_1, \ldots, X_n$ gilt

$$MSE = \mathrm{Var}\left[\frac{1}{n}\sum_{i=1}^{n} X_i\right] = \frac{1}{n^2}\sum_{i=1}^{n}\mathrm{Var}[X_i] = \frac{1}{n}\mathrm{Var}[X].$$

und damit $MSE \to 0$ für $n \to \infty$. $\square$

Schätzvariablen VI: Schwache Konsistenz

Definition 30

Ein Schätzer U eines Parameters θ ist schwach konsistent wenn

$$\lim_{n\to\infty} \Pr[|U - \theta| \geq \varepsilon] = 0$$

Satz 31

Ist ein Schätzer konsistent im quadratischen Mittel, dann ist er auch schwach konsistent. Insbesondere ist $\overline{X}$ *ein schwach konsistenter Schätzer für* X*.*

Beweis:

Mit der Ungleichung von Chebyshev gilt für ein beliebiges, festes $\varepsilon > 0$:

$$\Pr[|\overline{X} - \theta| \geq \varepsilon] = \Pr[|\overline{X} - \mathbb{E}[X]| \geq \varepsilon] \leq \frac{\mathrm{Var}[\overline{X}]}{\varepsilon^2} \to 0 \text{ für } n \to \infty$$

$\square$

Schätzvariablen VII: Schätzung der Varianz

Das Ergebnis einer wiederholten Messung wird präsentiert als

$$\overline{X} \pm S$$

wobei S^2 ein Schätzer für die Varianz von X darstellt.

Es liegt nah,

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

als Schätzer zu verwenden.

S_n^2 ist jedoch keinen erwartungstreuen Schätzer für die Varianz!

Schätzvariablen VIII: Schätzung der Varianz

Satz 32

Sei X *eine Zufallsvariable und sei*

$$S_{n-1}^2 := \frac{1}{n-1} \sum_{i=1}^{n} (X_i - \overline{X})^2.$$

Die Variable S_{n-1}^2 *ist ein erwartungstreuer Schätzer für* $\mathrm{Var}[X]$*.*

Schätzvariablen IX: Schätzung der Varianz

Beweis: Sei $\mu := \mathbb{E}[X] = \mathbb{E}[X_i] = \mathbb{E}[\overline{X}]$.

$$\begin{aligned}
(X_i - \overline{X})^2 &= (X_i - \mu + \mu - \overline{X})^2 \\
&= (X_i - \mu)^2 + (\mu - \overline{X})^2 + 2(X_i - \mu)(\mu - \overline{X}) \\
&= (X_i - \mu)^2 + (\mu - \overline{X})^2 - \frac{2}{n}\sum_{j=1}^{n}(X_i - \mu)(X_j - \mu) \\
&= \frac{n-2}{n}(X_i - \mu)^2 + (\mu - \overline{X})^2 - \frac{2}{n}\sum_{j\neq i}(X_i - \mu)(X_j - \mu).
\end{aligned}$$

Für je zwei unabhängige Zufallsvariablen X_i, X_j mit $i \neq j$ gilt

$$\mathbb{E}[(X_i - \mu)(X_j - \mu)] = \mathbb{E}[X_i - \mu] \cdot \mathbb{E}[X_j - \mu] \; = \; 0 \cdot 0 \; = \; 0.$$

Schätzvariablen X: Schätzung der Varianz

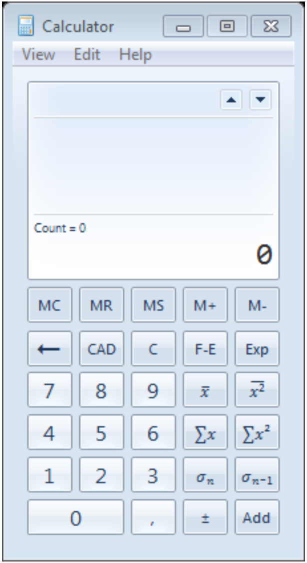
Daraus folgt

$$\begin{aligned}\mathbb{E}[(X_i - \overline{X})^2] &= \frac{n-2}{n} \cdot \mathbb{E}[(X_i - \mu)^2] + \mathbb{E}[(\mu - \overline{X})^2] \\ &= \frac{n-2}{n} \cdot \mathrm{Var}[X_i] + \mathrm{Var}[\overline{X}] \\ &= \frac{n-2}{n} \cdot \frac{1}{n} \cdot \mathrm{Var}[X] + \mathrm{Var}[\overline{X}] = \frac{n-1}{n} \cdot \mathrm{Var}[X]\end{aligned}$$

und somit gilt für S_{n-1}^2

$$\begin{aligned}\mathbb{E}[S_{n-1}^2] &= \frac{1}{n-1} \sum_{i=1}^{n} \mathbb{E}[(X_i - \overline{X})^2] \\ &= \frac{1}{n-1} \cdot n \cdot \frac{n-1}{n} \cdot \mathrm{Var}[X] = \mathrm{Var}[X].\end{aligned}$$

Schätzvariablen XI: Schätzung der Varianz



Schätzvariablen XII: Stichprobenmittel und -varianz

Definition 33

Die Zufallsvariablen

$$\overline{X} := \frac{1}{n}\sum_{i=1}^{n} X_i \quad \text{und} \quad S_{n-1}^2 := \frac{1}{n-1}\sum_{i=1}^{n}(X_i - \overline{X})^2$$

heißen Stichprobenmittel bzw. Stichprobenvarianz der Stichprobe $X_1, \ldots, X_n$.

$\overline{X}$ und S_{n-1}^2 sind erwartungstreue Schätzer für den Erwartungswert bzw. die Varianz von X.

Oft wird S^2 statt S_{n-1}^2 geschrieben.

Schätzvariablen XIII

Beispiel 34

Ein Program läuft 10 Mal mit demselben Input mit Laufzeiten (ms)

2582	2580	2568	2576	2582
2581	2623	2616	2622	2617

Damit gilt für (die Realisierungen von) $\overline{X}$ und S_{n-1}

$$\overline{X} \approx 2594,7 \qquad S_{n-1} \approx 21,8$$

und wir schreiben als Ergebnis: $2594.7 \pm 21.8.$

Unter den Annahmen der Normalverteilung beträgt die W'keit, dass die Laufzeit einer Ausführung ausserhalb des Intervalls

$$[2594.7 - 2 \cdot 21.8\ ,\ 2594.7 + 2 \cdot 21.8] = [2551.1\ ,\ 2638.3]$$

liegt, höchstens 5% .

21. Das Maximum-Likelihood-Prinzip

Maximum-Likelihood-Prinzip I

Das Maximum-Likelihood-Prinzip ist ein allgemeines Prinzip zur Bestimmung des Parameters einer W'eitsverteilung.

> *Wähle den Wert des Parameters, für den die beobachtete Stichprobe maximale W'keit hat.*

Beispiel 35

Folgendes Experiment wird 10 mal wiederholt: Eine Münze wird bis zum ersten „Kopf" geworfen.

Die Münze wird jeweils $4, 4, 7, 2, 4, 6, 5, 3, 5, 2$ Mal geworfen.

Frage: Welche ist die Erfolgsw'keit der Münze ?

Sei X die Anzahl der Würfe. Es gilt $X \sim Geo(p)$.

Nach dem Prinzip sollen wir p so wählen, das die W'keit einer Beobachtung $4, 4, 7, 2, 4, 6, 5, 3, 5, 2$ maximiert wird.

Maximum-Likelihood-Prinzip II: Formalisierung

Sei X eine Zufallsvariable mit $\Pr[X = x] = f(x; \theta)$ für eine bekannte Funktion f mit einem unbekannten Parameter θ.

Sei $X_1, \ldots, X_n$ Stichprobenvariablen mit Dichte $f(x; \theta)$.

Eine Stichprobe liefert für jede X_i einen Wert x_i und wir schreiben $\vec{x} = (x_1, \ldots, x_n)$.

Für eine feste Stichprobe $\vec{x}$ gibt die Likelihood-Funktion

$$L(\vec{x}; \theta) := \prod_{i=1}^{n} f(x_i; \theta) = \prod_{i=1}^{n} \Pr_\theta[X_i = x_i]$$

$$\stackrel{\text{unabh.}}{=} \Pr_\theta[X_1 = x_1, \ldots, X_n = x_n]$$

die Wahrscheinlichkeit an, dass man die Stichprobe $\vec{x}$ erhält, wenn der Parameter den Wert θ hat.

Maximum-Likelihood-Prinzip IV: Formalisierung

Definition 36

Ein Schätzwert $\widehat{\theta}$ für den Parameter einer Verteilung $f(x;\theta)$ heißt Maximum-Likelihood-Schätzwert (ML-Schätzwert) für eine Stichprobe $\vec{x}$, wenn gilt

$$L(\vec{x};\theta) \leq L(\vec{x};\widehat{\theta}) \text{ für alle } \theta.$$

Wenn $L(\vec{x};\theta)$ differenzierbar ist, dann kann ein Maximum von $L(\vec{x};\theta)$ mit Hilfe der üblichen Methode berechnet werden:

- Berechne $L'(\vec{x};\theta) := \frac{\mathrm{d}\, L(\vec{x};\theta)}{\mathrm{d}\, \theta}$.
- Finde eine Lösung θ_0 der Gleichung $L'(\vec{x};\theta) = 0$ mit $L''(\vec{x};\theta_0) > 0$.

Maximum-Likelihood-Prinzip V: Beispiele

Beispiel 37

Wir konstruieren mit der ML-Methode einen Schätzer für den Parameter p der Bernoulli-Verteilung.

Mit $\Pr_p[X_i = 1] = p$ und $\Pr_p[X_i = 0] = 1 - p$ gilt

$$\Pr_p[X_i = x_i] = p^{x_i}(1-p)^{1-x_i} \ .$$

Die Likelihood-Funktion ist

$$L(\vec{x}; p) = \prod_{i=1}^{n} p^{x_i} \cdot (1-p)^{1-x_i}$$

Maximum-Likelihood-Prinzip VI: Beispiele

Statt L maximieren wir $\ln L$ (einfachere Berechnung).

$$\begin{aligned}\ln L(\vec{x};p) &= \sum_{i=1}^{n}(x_i \cdot \ln p + (1-x_i)\cdot \ln(1-p)) \\ &= n\bar{x}\cdot \ln p + (n - n\bar{x})\cdot \ln(1-p).\end{aligned}$$

mit $\bar{x} = \dfrac{1}{n}\sum_{i=1}^{n} x_i$.

Nullsetzen der Ableitung ergibt:

$$\frac{\mathsf{d}\ln L(\vec{x};p)}{\mathsf{d}\,p} = \frac{n\bar{x}}{p} - \frac{n-n\bar{x}}{1-p} = 0.$$

mit Lösung $p = \bar{x}$, ein Maximum.

Maximum-Likelihood-Prinzip VII: Beispiele

Sei $X \sim \mathcal{N}(\mu, \sigma^2)$. Wir bestimmen Schätzer für μ und σ^2. Es gilt

$$L(\vec{x}; \mu, \sigma^2) = \left(\frac{1}{\sqrt{2\pi}\sigma}\right)^n \cdot \prod_{i=1}^{n} \exp\left(-\frac{(x_i - \mu)^2}{2\sigma^2}\right).$$

Durch Logarithmieren erhalten wir

$$\ln L(\vec{x}; \mu, \sigma^2) = -n(\ln\sqrt{2\pi} + \ln\sigma) + \sum_{i=1}^{n}\left(-\frac{(x_i - \mu)^2}{2\sigma^2}\right).$$

Maximum-Likelihood-Prinzip VII: Beispiele

Nullsetzen der Ableitungen ergibt

$$\frac{\partial \ln L}{\partial \mu} = \sum_{i=1}^{n} \frac{x_i - \mu}{\sigma^2} \stackrel{!}{=} 0$$

$$\frac{\partial \ln L}{\partial \sigma} = -\frac{n}{\sigma} + \sum_{i=1}^{n} \frac{(x_i - \mu)^2}{\sigma^3} \stackrel{!}{=} 0$$

und wir erhalten

$$\mu = \bar{x} \quad \text{und} \quad \sigma^2 = \frac{1}{n} \sum_{i=1}^{n} (x_i - \mu)^2.$$

Der zweiter Wert ist fast die Stichprobenvarianz S^2_{n-1} aber mit Vorfaktor $\frac{1}{n}$ statt $\frac{1}{n-1}$.

Die ML-Schätzvariable für die Varianz ist somit nicht erwartungstreu.

22. Konfidenzintervalle

Konfidenzintervalle I

Problem: Wie gut kann man sich auf einen Schätzwert verlassen?

Lösung: Berechne statt einem Schätzer U zwei Schätzer U_1 und U_2 mit

$$\Pr[U_1 \leq \theta \leq U_2] \geq 1 - \alpha.$$

Die W'keit $1 - \alpha$ heißt Konfidenzniveau und kann dem „Sicherheitsbedürfnis“ angepasst werden.

Informalle Bedeutung:

> *Wenn wir für eine Stichprobe die Schätzer U_1 und U_2 berechnen und $\theta \in [U_1, U_2]$ schliessen, dann irren wir höchstens mit W'keit α.*

$[U_1, U_2]$ heißt Konfidenzintervall.

Oft setzt man $U_1 := U - \delta$ und $U_2 := U + \delta$ für einen Schätzer U (symmetrisches Konfidenzintervall).

Konfidenzintervalle II: Normalverteilung

Beispiel 38

Sei $X \sim \mathcal{N}(\mu, \sigma^2)$ und seien $X_1, \ldots, X_n$ zugehörige Stichprobenvariablen.

Wir Schätzen μ durch $\overline{X}$.

Wir suchen ein symmetrisches Konfidenzintervall für $\overline{X}$, d.h. einen Wert δ mit

$$\Pr[\overline{X} - \delta \leq \theta \leq \overline{X} + \delta] \geq 1 - \alpha.$$

Es gilt (Satz 133) $\overline{X} \sim \mathcal{N}(\mu, \frac{\sigma^2}{n})$. Nach Lemma 127 ist

$$Z := \sqrt{n} \cdot \frac{\overline{X} - \mu}{\sigma}$$

standardnormalverteilt.

Konfidenzintervalle III: Normalverteilung

Für Z suchen wir nach einem Konfidenzintervall $[-c, c]$ mit

$$\Pr[-c \le Z \le c] \stackrel{!}{=} 1 - \alpha.$$

Auflösen nach μ ergibt

$$\Pr\left[\overline{X} - \frac{c\sigma}{\sqrt{n}} \le \mu \le \overline{X} + \frac{c\sigma}{\sqrt{n}}\right] \stackrel{!}{=} 1 - \alpha\,.$$

Für c muss also gelten:

$$\Pr[-c \le Z \le c] = \Phi(c) - \Phi(-c) \stackrel{!}{=} 1 - \alpha.$$

Mit $\Phi(-c) = 1 - \Phi(c)$ erhalten wir

$$2 \cdot \Phi(c) - 1 \stackrel{!}{=} 1 - \alpha$$

also

$$c = \Phi^{-1}\left(1 - \frac{\alpha}{2}\right).$$

Konfidenzintervalle IV: Normalverteilung

Definition 39

X sei eine stetige Zufallsvariable mit Verteilung F_X.
Die Zahl x_γ mit

$$F_X(x_\gamma) = \gamma$$

heißt γ-Quantil der Verteilung F_X.

Für die Standardnormalverteilung bezeichnet z_γ das γ-Quantil.

Damit kann das gesuchte Konfidenzintervall durch

$$\left[\, \overline{X} - \frac{z_{(1-\frac{\alpha}{2})}\sigma}{\sqrt{n}} \,,\, \overline{X} + \frac{z_{(1-\frac{\alpha}{2})}\sigma}{\sqrt{n}} \,\right]$$

angegeben werden.

23. Testen von Hypothesen

Testen von Hypothesen I: Einführung

Bislang haben wir versucht, Parameter von Verteilungen zu schätzen, z.B. die Erfolgsw'keit p einer Bernoulli-verteilte Variable. X.

In der Praxis ist man jedoch nur an einer Eigenschaft des Parameters interessiert, z.B. ob $p \geq 1/3$, $p \leq 0.8$, oder $p = 0.5$ gilt.

Statistische Tests werden verwendet, um mit einer gewissen Fehlerw'keit zu entscheiden, ob die Eigenschaft gilt oder nicht.

Testen von Hypothesen II: Terminologie

Definition 40

Die zu überprüfende Eigenschaft bezeichnet man mit H_0 und wird Nullhypothese genannt.

Man nimmt an, dass entweder die Nullhypothese oder die mit H_1 bezeichnete Alternative gilt.

Bei den meisten Tests gilt $H_1 := \neg H_0$ (triviale Alternative).

Der Test entscheidet, welche von den beiden Eigenschaften abgelehnt wird.

Ein Beispiel, in dem eine nicht triviale Alternative Sinn macht:

Beispiel 41

Von einer Festplatte ist bekannt, dass sie zu einer von zwei Baureihen gehört. Die mittleren Zugriffszeiten dieser Baureihen betragen 9ms bzw. 12ms.

Es soll mit einem statistischen Test entschieden werden, zu welchem Typ die betrachtete Festplatte gehört.

In diesem Test $H_0 : \mu \leq 9$ und $H_1 : \mu \geq 12$ mit μ die mittlere Zugriffszeit.

Testen von Hypothesen IV: Terminologie

Ein Test wiederholt n-Mal das zur Variablen X zugehörigen Zufallsexperiment. Seien $X_1, \dots, X_n$ unabhängige Stichprobenvariablen mit derselben Verteilung wie X.

Definition 42

Die Testgröße eines Tests ist eine Zufallsvariable der Gestalt $T := f(X_1, \dots, X_n)$.

Die Ablehnungsbedingung ist eine Bedingung der Gestalt

$$
\begin{array}{ll}
T < c \;\text{ oder }\; T > c & \text{(einseitiger Test)} \\
T < c_1 \;\vee\; T > c_2 & \text{(zweiseitiger Test)}
\end{array}
$$

Die Nullhypothese wird abgelehnt, wenn der Stichprobenvektor $\vec{x} = (x_1, \dots, x_n)$ die Ablehnungsbedingung erfüllt (d.h., $f(\vec{x}) < c$, $f(\vec{x}) > c$, oder $f(\vec{x}) < c_1 \vee f(\vec{x}) > c_2$ gilt).

Der Ablehnungsbereich ist die Menge der Stichprobenvektoren $\vec{x} = (x_1, \dots, x_n)$, für die die Nullhypothese abgelehnt wird.

Testen von Hypothesen V: Terminologie

Definition 43

Ein Fehler 1. Art tritt ein, wenn H_0 irrtümlich abgelehnt wird. (D.h., H_0 gilt aber die Stichprobe $\vec{x}$ liegt im Ablehnungsbereich.)

Ein Fehler 2. Art tritt ein, wenn H_0 irrtümlich angenommen wird. (D.h., H_0 gilt nicht aber die Stichprobe $\vec{x}$ liegt nicht im Ablehnungsbereich.)

Sei α eine obere Schranke der W'keit für den Fehler 1. Art. Wir sagen, dass der Test α-Fehler oder Signifikanzniveau von α hat.

In der Praxis gibt man sich einen Wert für α vor und bestimmt man den Ablehnungsbereich so, dass der Test Signifikanzniveu α hat.

Übliche Werte für α sind $0{,}05$, $0{,}01$ oder $0{,}001$.

Testen von Hypothesen VI: Terminologie

Die Minimierung des Fehlers 1. Art und des Fehlers 2. Art sind gegenläufige Ziele, z.B.:

- Lehnt man die Nullhypothese nie ab, hat ein Fehler 1. Art W'keit 0. Allerdings tritt ein Fehler 2. Art immer ein, wenn H_0 nicht gilt.
- Lehnt man die Nullhypothese immer ab, hat ein Fehler 2. Art W'keit 0. Allerdings tritt ein Fehler 1. Art immer ein, wenn H_0 gilt.

Ziel der meisten statistischen Tests ist eine kleine W'keit für den Fehler 1. Art zu garantieren. Die W'keit des Fehlers 2.Art kann groß sein!

Bei der Wahl der Nullhypothese (setzt man $H_0 = H$ oder $H_0 = \neg H$?) muss das berücksichtigt werden.

Testen von Hypothesen VII: Entwurf eines Tests

Wir entwerfen einen Test für den Parameter p einer Bernoulli-verteilten Zufallsvariablen X.

Wir setzen $H_0 : p \geq p_0$ und $H_1 : p < p_0$.

Als Testgröße verwenden wir $T := X_1 + \ldots + X_n$.
(Anzahl der Köpfe.)

Wir lehnen H_0 ab wenn der Wert von T "zu klein" ist.
Als Gestalt des Ablehnungsbereichs wählen wir also $T < k$.
(Einseitiger Test.)
Für eine Nullhypothese $H_0 : p = p_0$ würden wir $T < k_1 \ \vee \ T > k_2$ wählen.

Der Wert $k \in \mathbb{R}$ muss in Abhängigkeit des Signifikanzniveaus α festgelegt werden.

Testen von Hypothesen VIII: Entwurf eines Tests

Wir bestimmen k.

Es gilt $T \sim Bin(n,p)$. Bei großer Stichprobenumfang n ist die Variable

$$\tilde{T} := \frac{T - np}{\sqrt{np(1-p)}}.$$

annähernd standardnormalverteilt (siehe Korollar 14).

Wir berechnen für jeden Wert von k die maximale Fehlerw'keit über die möglichen Werten von p:

$$\begin{aligned}
\text{Fehlerw'keit 1. Art} &\leq \sup_{p \text{ erfüllt } H_0} \Pr_p[\text{„}H_0 \text{ wird abgelehnt“}] \\
&= \max_{p \geq p_0} \Pr_p[T < k] = \Pr_{p=p_0}[T < k] \\
\text{Fehlerw'keit 2. Art} &\leq \sup_{p \text{ erfüllt } H_1} \Pr_p[\text{„}H_1 \text{ wird abgelehnt“}] \\
&= \sup_{p < p_0} \Pr_p[T \geq k] = \Pr_{p=p_0}[T \geq k]
\end{aligned}$$

Testen von Hypothesen IX: Entwurf eines Tests

Für das gewünschte Signifikanzniveau α erhalten wir

$$\begin{aligned}\alpha &= \Pr_{p=p_0}[T < k]\\ &= \Pr_{p=p_0}\left[\tilde{T} < \frac{k-np}{\sqrt{np(1-p)}}\right]\\ &= \Pr\left[\tilde{T} < \frac{k-np_0}{\sqrt{np_0(1-p_0)}}\right] \approx \Phi\left(\frac{k-np_0}{\sqrt{np_0(1-p_0)}}\right).\end{aligned}$$

Testen von Hypothesen X: Entwurf eines Tests

Damit gilt: Für jedes

$$k \leq z_\alpha \sqrt{np_0(1-p_0)} + np_0$$

kann die Nullhypothese $p \geq p_0$ abgelehnt werden mit Signifikanzniveau α.

Das größte solche k minimiert den Fehler 2. Art.

Für diesen Wert gilt:

$$\text{Fehlerw'keit 1. Art} = \Pr_{p=p_0}[T < k]$$
$$\text{Fehlerw'keit 2. Art} = \Pr_{p=p_0}[T \geq k]$$

und damit $\text{Fehlerw'keit 2. Art} = 1 - \text{Fehlerw'keit 1. Art}$

Es ist also nicht möglich, beide W'keiten gleichzeitig zu reduzieren.

Testen von Hypothesen XI: Entwurf eines Tests

Beispiel 44

Ein Test soll mit Signifikanz $\alpha = 0.05$ bestimmen, ob die W'keit von Kopf bei einer Münze mindestens $1/3$ beträgt.

Wenn die Münze $n = 450$ Mal geworfen wird, dann wird die Hypothese $p \geq 1/3$ abgelehnt werde, wenn die Anzahl der Köpfe unter

$$\begin{aligned} k &= z_{0.05} \cdot \sqrt{450 \cdot \frac{1}{3} \cdot \frac{2}{3}} + 450 \cdot \frac{1}{3} \\ &= z_{0.05} \cdot 10 + 150 \\ &\approx -1.64 \cdot 10 + 150 = 133,6 \end{aligned}$$

liegt.

Testen von Hypothesen XI: Entwurf eines Tests

Bei echten Alternativtests werden für hinreichend große Stichproben beide Testfehler klein.

Beispiel 45

Bei zwei Sorten von gezinkten Münzen ist die Erfolgs'wkeit jeweils $p \geq 3/5$ und $p \leq 2/5$. Mit Hilfe eines Tests soll die Sorte einer gegebenen Münze bestimmt werden.

Unter dieser Annahme setzt man:

$$H_0 : p \geq 3/5 \qquad H_1 : p \leq 2/5.$$

Testen von Hypothesen XII: Entwurf eines Tests

$$\text{Fehlerw'keit 1. Art} \leq \sup_{p \text{ erfüllt } H_0} \Pr_p[\text{„}H_0 \text{ wird abgelehnt“}]$$
$$= \max_{p \geq 3/5} \Pr_p[T < k] = \Pr_{p=3/5}[T < k]$$
$$\approx \Phi\left(\frac{k - n \cdot 3/5}{\sqrt{n \cdot 6/25}}\right)$$

$$\text{Fehlerw'keit 2. Art} \leq \sup_{p \text{ erfüllt } H_1} \Pr_p[\text{„}H_1 \text{ wird abgelehnt“}]$$
$$= \sup_{p \leq 2/5} \Pr_p[T \geq k] = \Pr_{p=2/5}[T \geq k]$$
$$\approx 1 - \Phi\left(\frac{k - n \cdot 2/5}{\sqrt{n \cdot 6/25}}\right)$$

Testen von Hypothesen XIII: Entwurf eines Tests

Mit z.B. $n = 100$ und $\alpha = 0.05$ ergibt sich $k \approx 51.94$. Für diesen Wert erhalten wir

$$\begin{aligned}\text{Fehlerw'keit 2. Art} &\approx 1 - \Phi\left(\frac{51.94 - 40}{2\sqrt{6}}\right) \\ &\approx 1 - \Phi(2.437) \approx 1 - 0.9926 \\ &= 0.0074\end{aligned}$$

24. Praktische Anwendung statistischer Tests

Praktische Anwendung statistischer Tests

Das im vorhergehenden Abschnitt konstruierte Testverfahren taucht in der Literatur unter dem Namen approximativer Binomialtest auf.

Die folgende Tabelle 1 gibt einen Überblick über die Eckdaten dieses Tests.

Tabelle: **Approximativer Binomialtest**

Annahmen:

$X_1, \ldots, X_n$ seien unabhängig und identisch verteilt mit $\Pr[X_i = 1] = p$ und $\Pr[X_i = 0] = 1 - p$, wobei p unbekannt sei.
n sei hinreichend groß, so dass die Approximation aus Korollar 14 brauchbare Ergebnisse liefert.

Hypothesen:

$$\begin{array}{lllll} a) & H_0 : p = p_0 & \text{gegen} & H_1 : p \neq p_0, \\ b) & H_0 : p \geq p_0 & \text{gegen} & H_1 : p < p_0, \\ c) & H_0 : p \leq p_0 & \text{gegen} & H_1 : p > p_0. \end{array}$$

Testgröße:

$$Z := \frac{h - np_0}{\sqrt{np_0(1 - p_0)}},$$

wobei $h := X_1 + \ldots + X_n$ die Häufigkeit bezeichnet, mit der die Ereignisse $X_i = 1$ aufgetreten sind.

Ablehnungskriterium für H_0 bei Signifikanzniveau α:

$$\begin{array}{ll} a) & |Z| > z_{1-\alpha/2}, \\ b) & Z < z_{\alpha}, \\ c) & Z > z_{1-\alpha}. \end{array}$$

Allgemeines Vorgehen bei statistischen Tests

1. Schritt: Formulierung von Annahmen.
   Ganz ohne Annahmen kommt man meist nicht aus.
   Übliche Annahmen betreffen die Verteilung der Stichprobenvariablen und deren Unabhängigkeit.
2. Schritt: Formulierung der Nullhypothese.
3. Schritt: Auswahl des Testverfahrens.
4. Schritt: Durchführung des Tests und Entscheidung.

Wie findet man das richtige Testverfahren? I

Statistische Tests kann man nach mehreren Kriterien in Klassen einteilen.

- **Anzahl der beteiligten Zufallsgrößen**
  - Ein-Stichproben-Test.
    Es wird nur eine Zufallsgröße untersucht, für die eine Stichprobe erzeugt wird.
    Beispiel: Beträgt die mittlere Zugriffszeit auf einen Datenbankserver im Mittel höchstens 10ms?
  - Zwei-Stichproben-Test.
    Zwei Zufallsgrößen, für die jeweils eine Stichprobe erzeugt wird, werden verglichen.
    Beispiel: Hat Datenbankserver A eine kürzere mittlere Zugriffszeit als Datenbankserver B?

Wie findet man das richtige Testverfahren? II

- **Art der Nullhypothese**
  - Tests auf Lageparameter.
    Die Nullhypothese ist eine Aussage über Lageparameter der Verteilung wie den Erwartungswert oder die Varianz.
  - Tests auf eine vorgegebene Verteilung.
    Die Nullhypotese ist eine Aussage über den Verteilungstyp, z.B. dass die Zufallsgröße Normalverteilt ist.
  - Tests auf ein Maß für die Abhängigkeit verschiedener Zufallsgröße.
    Z.B. sagt die Nullhypothese, dass zwei Zufallsvariablen unabhängig sind.

Wie findet man das richtige Testverfahren? III

- **Annahmen über die Zufallsgrößen**
  - Bekannter/unbekannter Verteilungstyp.
  - Bekannter/unbekannter Erwartungswert.
  - Bekannte/unbekannte Varianz.

Ein-Stichproben-Tests für Lageparameter I: Gaußtest

Tabelle: **Gaußtest**

Annahmen:

$X_1, \ldots, X_n$ seien unabhängig und identisch verteilt mit $X_i \sim \mathcal{N}(\mu, \sigma^2)$, wobei σ^2 bekannt ist.
Alternativ gelte $\mathbb{E}[X_i] = \mu$ und $\operatorname{Var}[X_i] = \sigma^2$, und n sei groß genug.

Hypothesen:

$$\begin{array}{llll} a) & H_0 : \mu = \mu_0 & \textit{gegen} & H_1 : \mu \neq \mu_0, \\ b) & H_0 : \mu \geq \mu_0 & \textit{gegen} & H_1 : \mu < \mu_0, \\ c) & H_0 : \mu \leq \mu_0 & \textit{gegen} & H_1 : \mu > \mu_0. \end{array}$$

Testgröße:

$$Z := \frac{\overline{X} - \mu_0}{\sigma} \sqrt{n}\,.$$

Ablehnungskriterium für H_0 bei Signifikanzniveau α:

$$\begin{array}{ll} a) & |Z| > z_{1-\alpha/2}, \\ b) & Z < z_{\alpha}, \\ c) & Z > z_{1-\alpha}. \end{array}$$

Ein-Stichproben-Tests für Lageparameter II: Gaußtest

Der Gaußtest hat den Nachteil, dass man die Varianz σ^2 der Variablen $X_1, \ldots, X_n$ kennen muss.

Wenn diese unbekannt ist kann die Varianz durch die Stichprobenvarianz S^2 (siehe Definition 33) ersetzt werden.

Dies führt auf den so genannten t-Test.

Ein-Stichproben-Tests für Lageparameter III: t-Test

Tabelle: **t-Test**

Annahmen:

$X_1, \dots, X_n$ seien unabhängig und identisch verteilt mit $X_i \sim \mathcal{N}(\mu, \sigma^2)$.
Alternativ gelte $\mathbb{E}[X_i] = \mu$ und $\mathrm{Var}[X_i] = \sigma^2$, und n sei groß genug.

Hypothesen:

a) $H_0 : \mu = \mu_0$ *gegen* $H_1 : \mu \neq \mu_0$,
b) $H_0 : \mu \geq \mu_0$ *gegen* $H_1 : \mu < \mu_0$,
c) $H_0 : \mu \leq \mu_0$ *gegen* $H_1 : \mu > \mu_0$.

Testgröße:

$$T := \frac{\overline{X} - \mu_0}{S} \sqrt{n}.$$

Ablehnungskriterium für H_0 bei Signifikanzniveau α:

a) $|T| > t_{n-1,1-\alpha/2}$,
b) $T < t_{n-1,\alpha}$,
c) $T > t_{n-1,1-\alpha}$.

Ein-Stichproben-Tests für Lageparameter IV: t-Test

Hierbei bezeichnet $t_{n-1,1-\alpha}$ das

$(1-\alpha)$-*Quantil der t-Verteilung mit $n-1$ Freiheitsgraden*

an.

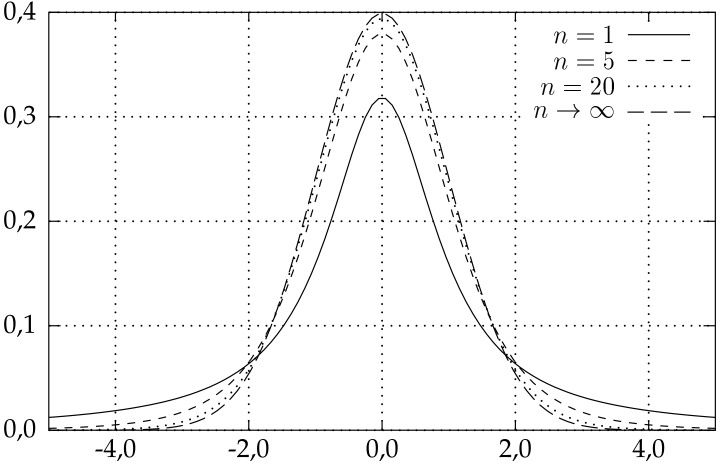
Die t-Verteilung nennt man auch Student-Verteilung.

Es handelt sich eigentlich um eine Familie von Verteilungen, eine für jede Anzahl von Freihaitsgraden.

Die Dichte ähnelt der Dichte der Normalverteilung.

Für große n (Faustregel: $n \geq 30$) wird in der Praxis die t-Verteilung durch die Normalverteilung angenähert.

Ein-Stichproben-Tests für Lageparameter V: t-Test



Dichte der t-Verteilung mit n Freiheitsgraden

Zwei-Stichproben-Tests für Lageparameter: t-Test

Tabelle: **Zwei-Stichproben-t-Test**

Annahmen:

$X_1, \ldots, X_m$ und $Y_1, \ldots, Y_n$ seien unabhängig und jeweils identisch verteilt, wobei $X_i \sim \mathcal{N}(\mu_X, \sigma_X^2)$ und $Y_i \sim \mathcal{N}(\mu_Y, \sigma_Y^2)$ gelte.
Die Varianzen seien identisch, also $\sigma_X^2 = \sigma_Y^2$.

Hypothesen:

$$
\begin{array}{llcl}
\textit{a)} & H_0 : \mu_X = \mu_Y & \textit{gegen} & H_1 : \mu_X \neq \mu_Y\,, \\
\textit{b)} & H_0 : \mu_X \geq \mu_Y & \textit{gegen} & H_1 : \mu_X < \mu_Y\,, \\
\textit{c)} & H_0 : \mu_X \leq \mu_Y & \textit{gegen} & H_1 : \mu_X > \mu_Y\,.
\end{array}
$$

Testgröße:

$$T := \sqrt{\frac{n+m-2}{\frac{1}{m}+\frac{1}{n}}} \cdot \frac{\overline{X}-\overline{Y}}{\sqrt{(m-1)\cdot S_X^2 + (n-1)\cdot S_Y^2}}.$$

Ablehnungskriterium für H_0 bei Signifikanzniveau α:

$$
\begin{array}{ll}
\textit{a)} & |T| > t_{m+n-2,1-\alpha/2}\,, \\
\textit{b)} & T < t_{m+n-2,\alpha}\,, \\
\textit{c)} & T > t_{m+n-2,1-\alpha}\,.
\end{array}
$$

Tests auf Verteilungen I: χ^2-Anpassungstest

Tabelle: **χ^2-Anpassungstest**

Annahmen:

$X_1, \ldots, X_n$ *seien unabhängig und identisch verteilt mit* $W_{X_i} = \{1, \ldots, k\}$.

Hypothesen:

$$H_0 \;:\; \Pr[X = i] = p_i \quad \text{für } i = 1, \ldots, k,$$
$$H_1 \;:\; \Pr[X = i] \neq p_i \quad \text{für mindestens ein } i \in \{1, \ldots, k\},$$

Testgröße:

$$T = \sum_{i=1}^{k} \frac{(h_i - np_i)^2}{np_i},$$

wobei h_i *die Häufigkeit angibt, mit der* $X_1, \ldots, X_n$ *den Wert* i *angenommen haben.*

Ablehnungskriterium für H_0 bei Signifikanzniveau α:

$$T > \chi^2_{k-1,1-\alpha};$$

dabei sollte gelten, dass $np_i \geq 1$ *für alle* i *und* $np_i \geq 5$ *für mindestens* 80% *der Werte* $i = 1, \ldots, k$.

Tests auf Verteilungen II: χ^2-Anpassungstest

Für die Testgröße T wird näherungsweise eine

χ^2-Verteilung mit $k-1$ Freiheitsgraden

angenommen.

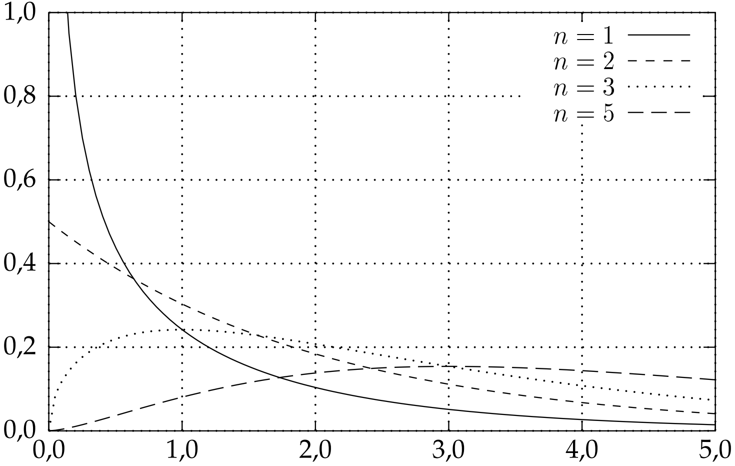
Das γ-Quantil einer χ^2-Verteilung mit k Freiheitsgraden bezeichen wir mit $\chi^2_{k,\gamma}$.

Die Werte dieser Verteilung finden sich in entsprechenden Tabellen in der Literatur.

Damit die Approximation gerechtfertigt ist, sollte gelten (Faustregel)

- $np_i \geq 1$ für alle i und
- $np_i \geq 5$ für mindestens 80% der Werte $i = 1, \ldots, k$.

Tests auf Verteilungen III: χ^2-Anpassungstest



Dichte der χ^2-Verteilung mit n Freiheitsgraden

Tests auf Verteilungen IV: χ^2-Anpassungstest

Beispiel 46

Wir wollen überprüfen, ob der Zufallszahlengenerator von Maple eine gute Approximation der Gleichverteilung liefert.

Dazu lassen wir Maple $n = 100000$ Zufallszahlen aus der Menge $\{1, \ldots, 10\}$ generieren.

Die Nullhypothese lautet $p_1 = \ldots = p_{10} = 1/10$, wobei p_i die W'keit von i bezeichnet.

Die Nullhypothese soll mit Signifikanzniveau von $\alpha = 0{,}05$ getestet werden.

Tests auf Verteilungen V: χ^2-Anpassungstest

Beispiel:

i	1	2	3	4	5	6	7	8	9	10
h_i	10102	10070	9972	9803	10002	10065	10133	9943	10009	9901

Es gilt $T = 8{,}9946$ und $\chi^2_{9,0{,}95} \approx 16{,}919$.

Die Nullhypothese wird nicht abgelehnt.

Was bedeutet das?

Nicht viel! Der Test garantiert nur, dass die W'keit, die Nullhypothese irrtümlich abzulehnen, höchstens 5% beträgt. Die Nullhypothese ist jedoch nicht abgelehnt worden!