3 Statistische Tests von Pseudozufallszahlen

Def. 4.1 *Ein* <u>*Test*</u> *ist eine Entscheidungsvorschrift, die über die Akzeptanz genau einer von zwei alternativen Hypothesen entscheidet.*

Sehen wir uns das Prinzip von statistischen Tests anhand eines Beispiels an:

Ein Käufer soll mit Hilfe einer stichpunktartigen Qualitätskontrolle entscheiden, ob er einen Warenbestand kauft oder nicht. Es gibt dabei zwei Hypothesen, die Nullhypothese und die Alternativhypothese:

H_0 **:** Die Ware ist in Ordnung,
z.B. der Ausschußanteil p ist kleiner oder gleich 2%.

H_a**:** Die Ware ist schlecht,
d.h. $p > 2\%$.

Der Kunde führt nun bei n Produkten eine Kontrolle durch (Stichproben) und bewertet das jeweilige Ergebnis seiner Probe durch die 'Beobachtungen' $x_1, \ldots, x_n$, wobei:

$$x_i = \begin{cases} 0 & \text{, falls das Produkt } i \text{ gut ist,} \\ 1 & \text{, falls das Produkt } i \text{ schlecht ist.} \end{cases}$$

Dann ist $z = \sum_{i=1}^{n} x_i$ die Anzahl der fehlerhaften Produkte, die der Kunde gefunden hat. Nun wird vor dem Test ein kritischer Wert z_α festgelegt

- Ist $z > z_\alpha$, so wird die Hypothese H_0 abgelehnt;
- Ist $z \leq z_\alpha$, so wird die Hypothese H_0 für richtig befunden.

In diesem Zusammenhang sind zwei (bedingte) Wahrscheinlichkeiten wichtig:

1. $P(Z > z_\alpha | H \text{ ist wahr})$ – die Wahrscheinlichkeit also, daß der Käufer die Ware für schlecht befindet und ablehnt, obwohl sie doch in Ordnung ist. Diese Wahrscheinlichkeit spiegelt das „Risiko des Produzenten“ wider.

2. $P(Z \leq z_\alpha | H \text{ ist falsch})$ – die Wahrscheinlichkeit also, daß der Käufer die Ware nimmt, obwohl ihre Qualität stark zu wünschen übrig läßt. Diese Wahrscheinlichkeit spiegelt das „Risiko des Käufers“ wider.

Die Entscheidung für H_A oder für H_0 wird anhand einer Teststatistik

$$Z = Z(x_1, ..., x_n)$$

gefällt. Zeigt der Wert von Z in einem vorher bestimmten Bereich K, dem sogen. Ablehnungsbereich oder kritischen Bereich, dann wird H_0 abgelehnt, anderenfalls wird H_0 nicht abgelehnt.

Bei jeder dieser Entscheidungen kann man Fehlentscheidungen treffen:

Entscheidung für H_A obwohl H_0 richtig ist:

Fehler 1.Art

Entscheidung für H_0 obwohl H_A richtig ist:

Fehler 2.Art

	Entscheidung für H_0	Entscheidung für H_A
H_0 richtig	richtig, Sicherheitswkt. $1-\alpha$	Fehler 1. Art Fehlerwkt. α.
H_A richtig	Fehler 2.Art Fehlerwkt. 1-β	richtig, Güte β

Bem.: Entscheidung für H_0 heißt nicht notwendig, dass H_0 richtig ist.

Der Parameter α wird dabei auf $P(Z > Z_\alpha | H$ ist wahr$)$ gesetzt und meist vorgegeben. Übliche Werte für α sind 0,01 oder 0,05. Gesucht ist eine Testvorschrift, die zur Minimierung des „Risikos des Käufers“ führt.

Anwendung auf Pseudozufallszahlen

zu testen:

- Gleichverteilung der Pseudozufallszahlen über dem Intervall $[0, 1[$;
- Unabhängigkeit der Pseudozufallszahlen.

Angen., unsere Pseudozufallszahlen erfüllten diese beiden Eigenschaften. Dann gilt für $0 \leq a < b \leq 1$:

$$P(a \leq u_i \leq b) = b - a.$$

Weiterhin gilt:

$$\frac{\#\{u_i \in \{u_1, ..., u_n\} : \ a \leq u_i \leq b\}}{n} \xrightarrow[n \to \infty]{} P(a \leq u_i \leq b) = b - a. \quad (1)$$

Außerdem gilt:

$$F_n(x) = \frac{\#\{u_i \in \{u_1, ..., u_n\} : \ u_i < x, \, 0 \leq x < 1\}}{n} \approx x = F(x) \quad (2)$$

(Satz von Glivenko).

3.1 Test auf Gleichverteilung

Der χ^2–Anpassungs-Test

Def. 4.2 *$X_1, \ldots, X_k$ seien unabhängig, identisch verteilte Zufallszahlen mit $X_i \sim N(0,1)$.*

Dann heißt die Zufallszahl Y mit

$$Y = \sum_{i=1}^{k} X_i^2$$

χ^2-verteilt mit k Freiheitsgraden.

Bez. 7 $Y \sim \chi_k^2$.

Es seien X_i $(i = 1, \ldots, n)$ beliebige unabhängig und identisch verteilte Zufallsgrößen und $A_j (j = 1, \ldots, k)$ zufällige Ereignisse (Intervalle in $\mathbb{R}$).

$$P(X_i \in B) = 1, \quad B = \bigcup_{j=1}^{k} A_j, \quad A_j \cap A_i = \emptyset \quad \forall i \neq j$$

$$P(X_i \in A_j) = p_j$$

Wir definieren

$$\chi^2 = \sum_{j=1}^{k} \frac{(n_j - n \cdot p_j)^2}{np_j} \qquad n_j = \#\{X_i : X_i \in A_j\}$$

Für große n gilt dann approximativ, $\chi^2 \sim \chi^2_{k-1}$.

Wir testen

$$\begin{aligned} H_0 : &\ P(X_i \in A_j) = p_j \quad j = 1, \ldots, k \\ H_A : &\qquad\qquad P(X_i \in A_j) \neq p_j. \end{aligned}$$

Wenn H_0 richtig ist, gilt wegen dem schwachen Gesetz großer Zahlen (siehe Seite 296): $n_j \approx n \cdot p_j$

Offenbar, $0 \leq \chi^2$.

Wenn $\chi^2 \leq c_\alpha$ wollen wir Hypothese H_0 annehmen, wenn $\chi^2 > c_\alpha$ lehnen wir diese ab.

c_α wird wie folgt festgelegt:

$$P(\chi^2 > c_\alpha | H_0 \text{ richtig}) = \alpha$$

ist die Wahrscheinlichkeit (bzw. das Risiko) dafür, das trotz “guter” Verteilung (Gleichverteilung) der Zufallszahlen wir die Hypothese H_0 ablehnen, d.h. die Nicht-Gleichverteilung annehmen.

`ZufallszahlenMuster.sas`

Anwendung auf Zufallszahlen

$$\begin{aligned}
B &= [0,1) \\
A_j &= \left[\frac{j-1}{k}, \frac{j}{k}\right), \qquad n \geq 5k \\
p_j &= \frac{1}{k} \\
\chi^2 &= \sum_{j=1}^{k} \frac{(n_j - \frac{n}{k})^2}{\frac{n}{k}} \\
&\sim \chi^2_{k-1}, \quad \text{falls } p_j = \frac{1}{k} \quad \forall j = 1, 2, \ldots, k
\end{aligned}$$

Auf der empirischen Verteilungsfunktion beruhende Tests (allgemein)

Def. (empirische Verteilungsfunktion):

Seien $X_1, ..., X_n$ unabh. Beobachtungen,

$X_{(1)} \leq ... \leq X_{(n)}$ die geordneten Beob.

Die Funktion

$$F_n(x) = \begin{cases} 0 & x < X_{(1)} \\ \frac{i}{n} & X_{(i)} \leq x < X_{(i+1)} \quad i = 1...n \\ 1 & X_{(n)} \leq x \end{cases}$$

heißt empirische Verteilungsfunktion.

Satz v. Glivento-Cantelli: $F_n(x) \to F(x)$.

```
EDF.sas
EDF_2.sas
```

*Bitte keinen Schreck bekommen, die folgenden beiden Folien sind nur für den interessierten Leser.

Kolmogorov-Smirnov-Test

$$\begin{aligned} D &= \sup_x |F_n(x) - F_0(x)| \\ &= \max\Big(\max_i (\frac{i}{n} - U_{(i)}), \max_i (U_{(i)} - \frac{i-1}{n}) \Big) \end{aligned}$$

Anderson-Darling-Test

$$\begin{aligned} A^2 &= n \int_{-\infty}^{\infty} \frac{(F_n(x) - F_0(x))^2}{F_0(x)(1 - F_0(x))} dF_0(x) \\ &= -n - \frac{1}{n} \sum_{i=1}^{n} (2i-1) \big(\ln U_{(i)} + \ln(1 - U_{(n+1-i)}) \big) \end{aligned}$$

Cramer-von Mises-Test

$$\begin{aligned} W^2 &= n \int_{-\infty}^{\infty} \big(F_n(x) - F_0(x)\big)^2 dF_0(x) \\ &= \frac{1}{12n} + \sum_{i=1}^{n} \big(U_{(i)} - \frac{2i-1}{2n}\big)^2 \end{aligned}$$

$U_{(i)} = F_0(X_{(i)}) \qquad X_{(1)} \leq \ldots X_{(n)}$

hier: $F_0(x) = x$.

Modifikationen für endliche Stichproben

D**:** $D \cdot (\sqrt{n} - 0.01 + 0.85/\sqrt{n})$

A^2**:** $AD^2 \cdot (1.0 + 0.75/n + 2.25/n^2)$

W^2**:** $CM^2 \cdot (1.0 + 0.5/n)$

Kritische Werte

W^2: D'Agostino, Stephens (1986), S. 123.

A^2: Crawford Moss u.a. (1990)

`Test_GoF_Banknote.sas`

`Test_GoFDarwin.sas`

`aufg19.sas`

Der Kolmogoroff–Smirnov–Test

Erinnerung:

$$\lim_{n\to\infty} D_n = \lim_{n\to\infty} \sup_x |F_n(x) - x| = 0$$

Satz 4.3 (Kolmogoroff–Smirnov**)** *Es gilt:*

$$\lim_{n\to\infty} P(\sqrt{n} \cdot D_n < x) = \begin{cases} 1 + 2\sum_{i=1}^{\infty}(-1)^i \cdot e^{-2\cdot i^2\cdot x^2}, & x > 0, \\ 0, & \textit{sonst.} \end{cases}$$
$$=: Q(x)$$

Bem. 21 $Q(x)$ *ist die Verteilungsfunktion der Kolmogorov-Verteilung (Kolmogorov, 1933).*

Praktische Durchführung

1. Die Pseudozufallszahlen werden der Größe nach geordnet. Wir ermitteln also eine Permutation $u_{(1)}, \ldots, u_{(n)}$ der Zahlen, für die gilt:

$$u_{(1)} < u_{(2)} < \ldots < u_{(n)}.$$

(Voraussetzung ist dabei natürlich, daß die Periode des verwendeten Zufallszahlengenerators nicht kleiner ist als n, d.h. daß alle Zahlen verschieden voneinander sind.)

2. Wir bestimmen die empirische Verteilungsfunktion $F_n(x) = \frac{\#\{u_i:\ u_i<x,\, 0\leq x<1\}}{n}$.

3. Wir ermitteln die Zahl $D_n := \sup\limits_x |F_n(x) - x| = \max\left\{\max\limits_{1\leq i\leq n} a_i, \max\limits_{1\leq i\leq n} b_i\right\}$, wobei für alle $i = 1, \ldots, n$ gilt:

$$a_i := \left|u_{(i)} - \frac{i}{n}\right|, \quad b_i := \left|u_{(i)} - \frac{i-1}{n}\right|.$$

4. Mit einem vorgegebenen Schwellwert c_α führen wir nun folgenden Test durch:

$$\sqrt{n}\cdot D_n > c_\alpha \implies \text{Ablehnung der Hypothese } H$$

$$\sqrt{n}\cdot D_n \leq c_\alpha \implies \text{Annahme der Hypothese } H$$

Dabei ist

$$\alpha = P(H \text{ abgelehnt}|H_0) = P(\sqrt{n}\cdot D_n > c_\alpha|H_0).$$

Daraus folgt, daß gilt:

$$Q(c_\alpha) = \lim_{n\to\infty} P(\sqrt{n} \cdot D_n < c_\alpha) = 1 - \alpha.$$

Die folgende Tabelle zeigt für vorgegebene Parameter α den jeweiligen Wert von c_α:

α	c_α (gerundet)
0.01	1.63
0.05	1.36
0.1	1.22

3.2 Test auf Unabhängigkeit

Der Run–Test

Def. 4.3 *Jeder Teilabschnitt einer Folge unabhängiger, identisch verteilter Zufallszahlen, in dem die Zufallszahlen in aufsteigend geordnet sind, heißt* <u>*Run*</u>*.*

Bsp. 4.5 *Wir teilen eine Folge in Runs ein:*

Folge von Zufallszahlen	2	1 2 3	2 4	1 7 8 9	0
Run	*I.*	*II.*	*III.*	*IV.*	*V.*
Länge des Runs	1	3	2	4	1

Satz 4.4 *Es sei* $u_1, \ldots, u_n$ *eine Folge unabhängiger Zufallsgrößen mit* $u_i \sim \mathrm{U}(0,1)$ *(*$i = 1, \ldots, n$*). Dann gilt für die Länge* R *eines Runs:*

$$P(R = r) = \frac{r}{(r+1)!}.$$

Wir fassen also R als Zufallsgröße auf:

$$R: \quad \begin{pmatrix} 1 & 2 & \ldots & r & \ldots \\ \frac{1}{2} & \frac{1}{3} & \ldots & \frac{r}{(r+1)!} & \ldots \end{pmatrix}.$$

Beweis: Wegen der Unabhängigkeit und der identischen Verteilung genügt es, die ersten $r+1$ Zufallsvariablen zu betrachten. Es gilt:

$$\begin{aligned} P(R=r) &= P(U_1 \le \cdots \le U_r > U_{r+1}) \\ &= P(U_1 \le \cdots \le U_r) - P(U_1 \le \cdots \le U_r \le U_{r+1}) \\ &= \frac{1}{r!} - \frac{1}{(r+1)!} = \frac{r}{(r+1)!} \end{aligned}$$

□

Bem. 22 *Es gilt:*

$$\begin{aligned} \sum_{i=1}^{\infty} P(R=i) &= \frac{1}{i!} - \frac{1}{(i+1)!} \\ &= \sum_{i=1}^{\infty} \frac{1}{i!} - \sum_{i=1}^{\infty} \frac{1}{(i+1)!} \\ &= (\sum_{i=0}^{\infty} \frac{1}{i!} - 1) - (\sum_{i=0}^{\infty} \frac{1}{(i+1)!} - 1) \\ &= (e-1) - (e-1-1) = 1. \end{aligned}$$

Es seien nun $u_1, \ldots, u_n$ Pseudozufallszahlen. Wir testen

$$H_0: \quad u_1, \ldots, u_n \quad \text{sind unabhängig} \qquad \text{gegen}$$
$$H_1: \quad u_1, \ldots, u_n \quad \text{sind abhängig.}$$

$R_1, \ldots, R_m$ sei die Folge der Längen der auftretenden Runs. Da Pseudozufallszahlen durch einen deterministischen Algorithmus berechnet werden, sind diese Längen auf keinen Fall unabhängig. Für unseren Test ist die Unabhängigkeit aber dringend notwendig. Deshalb streichen wir nach jedem Run die nächste Zufallszahl. Es entstehen die Größen $R_1^*, \ldots, R_m^*$. Formal sieht das folgendermaßen aus:

Wir bilden zunächst die folgenden Infima:

$$\begin{aligned}
S_1 &= \inf\{n \in \mathbb{N}\colon u_{n+1} < u_n\} \\
S_2 &= \inf\{n \in \mathbb{N}\colon n > S_1 + 1, u_{n+1} < u_n\} \\
&\vdots \\
S_{k+1} &= \inf\{n \in \mathbb{N}\colon n > S_k + 1, u_{n+1} < u_n\}
\end{aligned}$$

Dann definieren wir:

$$\begin{aligned} R_1^* &:= S_1 \\ R_2^* &:= S_2 - S_1 - 1 \\ &\vdots \\ R_{k+1}^* &:= S_{k+1} - S_k - 1 \end{aligned}$$

Wenn nun die Hypothese H_0 gilt, dann ist:

$$P(R^* = r) = \frac{r}{(r+1)!},$$

und die R_i^* ($i = 1, \ldots, m$) sind unabhängig.

Run-Test: Anpassungstest auf diese Verteilung

Teilen $\mathbf{Z}^+$ in disjunkte Teilintervalle auf:

$$[i_1 + 1, i_2], [i_2 + 1, i_3], \ldots, [i_k + 1, \infty)$$

$$0 = i_1 < i_2 < \ldots < i_k < \infty \qquad k \text{ Intervalle}$$

$$p_j^* = \sum_{l=i_j+1}^{i_{j+1}} P(R^* = l) = P(i_j + 1 \le R^* \le i_{j+1})$$

$$n_j = \#_{i=1,\ldots,m}\{R_i^* : i_j + 1 \le R_i^* \le i_{j+1}\}$$

$$\chi^2 = \sum_{j=1}^{k} \frac{(n_j - mp_j^*)^2}{mp_j^*} \sim \chi^2_{k-1}$$

Falls H_0 richtig ist, d.h. unabhängige, gleichverteilte Zufallszahlen liegen vor, dann gilt:

$$n_j \approx mp_j^*.$$

Falls $\chi^2 >$ kritischer Wert, lehnen wir dir Unabhängigkeitshypothese ab.

Bem. 23 *Gesamtumfang der zu erzeugenden Zufallszahlen sollte* ≥ 4000 *sein.*
Wir haben hier einen Anpassungstest auf eine gegbene diskrete Verteilung gemacht.
χ^2*-Anpassungstests (auf eine stetige Verteilung, hier Gleichverteilung) sollten, u.a. wegen der Willkür der Klasseneinteilung mit Vorsicht betrachtet werden.*

Autokorrelationstest

Sei $u_1, \ldots, u_n$ eine Folge von zufälligen Variablen. Für alle m können wir nun bilden:

$$\rho_m(k) = \frac{\operatorname{cov}\left(u_m, u_{m+k}\right)}{\sigma_{u_m}\sigma_{u_{m+k}}}$$

wobei $1 \leq k \leq \frac{n}{2}$ und

$$u_{m+k} := u_{m+k-n}, \text{ falls } m + k > n$$

Wenn $u_1, \ldots, u_n$ identisch verteilt so $\sigma_{u_j} = \sigma \quad \forall j$ und

$$\operatorname{cov}\left(u_m, u_{m+k}\right) = \operatorname{cov}\left(u_1, u_{k+1}\right)$$

Autokorrelation k-ter Ordnung

$$\sigma_m(k) = \rho(k) = \frac{E(u_m \cdot u_{m+k}) - (Eu_m)^2}{\sigma^2}$$

$\forall m, \quad k = 1, \ldots, \left[\frac{n}{2}\right]$

$$\hat{\rho}(k) = \frac{\frac{1}{n}\sum_{i=1}^n u_i \cdot u_{i+k} - \left(\sum_{i=1}^n u_i\right)^2}{\frac{1}{n}\sum_{i=1}^n u_i^2 - \left(\sum_{i=1}^n u_i\right)^2}$$

ist die *empirische Autokorrelation* k-ter Ordnung.

Falls nun $u_1, \ldots, u_n$ echte Zufallszahlen, dann gilt:

$$\rho(k) = 0 \quad \forall k \geq 1 \qquad\qquad \hat{\rho}(k) \approx 0$$

Bem. 24 *$\rho(k)$ ist die Pearson-Korrelation zwischen zwischen U_i und U_{i+k}.*

Ersetzen wir die

U_i durch ihre Ränge $R_1, \ldots, R_n$ und die

U_{i+k} durch ihre Ränge $S_1, \ldots, S_n$

dann erhalten wir den Spearman-Rang-Korrelationskoeffizient r_S. Es gilt asymptotisch

$$r_S \sim \mathcal{N}(0, \frac{1}{n-1}).$$

Die Nullhypothese

H_0: keine Autokorrelation

wird also abgelehnt, wenn

$$\sqrt{n-1}|r_S| \geq z_{1-\alpha/2}$$

($z_{1-\alpha/2}$: $1-\alpha/2$-Quantil der Standard-Normalverteilung,

$z_{0.975} = 1.96$.