# 5.8 Anpassungstests


8.1 Einführung
empirische Verteilungsfunktion

8.2 EDF-Anpassungstests
Kolmogorov-Smirnov-Test
Anderson-Darling-Test
Cramer-von Mises-Test

8.3 Anpassungstest auf Normalverteilung -
Shapiro-Wilk-Test

8.4. Anpassungstests auf weitere Verteilungen

# Anpassungstests

## Einführung

**Problem**

Klassische Test- und Schätzverfahren sind oft konzipiert unter der Normalverteilungsannahme.

**Frage**

Gilt sie überhaupt?

# Gilt die Normalverteilung? (1)

**Hampel, 1980**, Biometr. Journal

Eine Zeitlang glaubte (fast) jeder an das ’normale Fehlergesetz’,

die Mathematiker, weil sie es für ein empirisches Faktum hielten,

und die Anwender, weil sie es für ein mathematisches Gesetz hielten.

# Gilt die Normalverteilung? (2)

**Geary 1947**, Biometrika

Normality is a myth;

there never was,

and never will be,
a normal distribution.

# Anpassungstests

$(X_1, \ldots, X_n)$ iid., $X_i \sim F$, $F$ unbekannt.

**Anpassungstest auf eine spezifizierte Verteilung:**

$$H_0 : \quad F = F_0 \qquad \text{gegen} \qquad H_1 : \quad F \neq F_0.$$

I.A. hängt $F$ von unbekannten Parametern ab.

**Anpassungstest auf eine Normalverteilung:**

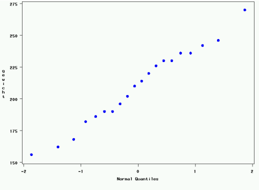
$$\begin{aligned} H_0 : F(x) &= \Phi\Big(\frac{x-\mu}{\sigma}\Big) \quad (\mu, \sigma \text{ unbekannt}) \\ H_1 : F(x) &\neq \Phi\Big(\frac{x-\mu}{\sigma}\Big) \quad \forall \mu, \sigma, \sigma > 0 \end{aligned}$$

($\Phi$: Verteilungsfunktion der Standardnormal.).

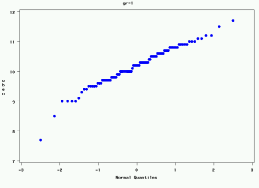
# Anpassungstests

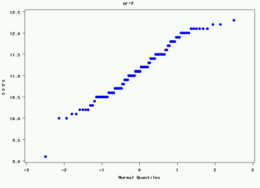
## Gewicht von Hühnern

| | | | | |
|---|---|---|---|---|
| 156 | 162 | 168 | 182 | 186 |
| 190 | 190 | 196 | 202 | 210 |
| 214 | 220 | 226 | 230 | 230 |
| 236 | 236 | 242 | 246 | 270 |




## Abmessungen von Banknoten, oben (echt, falsch)



# Auf der empirischen Verteilungsfunktion beruhende Tests

## Empirische Verteilungsfunktion

Seien $X_1, ..., X_n$ unabh. Beobachtungen,
$X_{(1)} \leq ... \leq X_{(n)}$ die geordneten Beob.
Die Funktion

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

heißt empirische Verteilungsfunktion.

Satz v. Glivento-Cantelli: $F_n(x) \rightarrow F(x)$.
(Hauptsatz der math. Statistik genannt)

# Anpassungstests

## Auf der empirischen Verteilungsfunktion beruhende Tests

### Kolmogorov-Smirnov-Test

$$D = \sup_x \left| F_n(x) - F_0(x) \right|$$

### Anderson-Darling-Test

$$\text{A-sq} = n \int_{-\infty}^{\infty} \frac{(F_n(x) - F_0(x))^2}{F_0(x)(1 - F_0(x))} dF_0(x)$$

### Cramer-von Mises-Test

$$\text{W-sq} = n \int_{-\infty}^{\infty} \big(F_n(x) - F_0(x)\big)^2 dF_0(x)$$

# Anpassungstests auf Normalverteilung

## Auf der empirischen Verteilungsfunktion beruhende Tests

hier:

$$\begin{aligned} F_0(x) &= \Phi\Big(\frac{x-\hat{\mu}}{\hat{\sigma}}\Big), \\ \hat{\mu} = \overline{X} &= \frac{1}{n}\sum_{i=1}^{n} X_i, \\ \hat{\sigma}^2 = s^2 &= \frac{1}{n-1}\sum_{i=1}^{n}(X_i - \overline{X})^2 \end{aligned}$$

$D \sim D_n$ (Kolmogorov-Verteilung) approx.

$$\lim_{n\to\infty} P_0\Big(D < \frac{x}{\sqrt{n}}\Big) = 1 - 2\sum_{k=1}^{\infty}(-1)^k e^{-2k^2x^2}$$

(Kolmogorov, 1933).

# Anpassungstests auf Normalverteilung

## Auf der empirischen Verteilungsfunktion beruhende Tests

Modifikationen für endliche Stichproben (zur Info.)

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

$\mathbf{A-sq}$: $A$-sq $\cdot(1.0 + 0.75/n + 2.25/n^2)$

**W**-sq: $W$-sq $\cdot(1.0 + 0.5/n)$

Große Werte von $D$, A-sq und W-sq führen jeweils zur Ablehnung von $H_0$.
$p$-Werte werden vom Programm berechnet.

```
Test_GoF_Banknote.sas
Test_GoFDarwin.sas
aufg24.sas
```

# Anpassungstests

## Shapiro-Wilk-Test

Vorbemerkungen:

$$X_i \sim \mathcal{N}(\mu, \sigma^2), \qquad Y_i = \frac{X_i - \mu}{\sigma} \sim \mathcal{N}(0, 1)$$

$i = 1, \ldots, n$.
Geordnete Beobachtungen:

$$X_{(1)} \leq \ldots \leq X_{(n)} \qquad Y_{(1)} \leq \ldots \leq Y_{(n)}.$$

Die Erwartungswerte

$$\begin{aligned} m_i &:= E(Y_{(i)}) \\ &= \frac{n!}{(i-1)!(n-i)!} \cdot \int_{-\infty}^{\infty} t\Phi^{i-1}(t)(1-\Phi(t))^{n-i}\phi(t)dt \end{aligned}$$

sind bekannt (und vertafelt).

# Shapiro-Wilk-Test

Approximation (Blom, 1958)

$$m_i \approx \tilde{m}_i = \Phi^{-1}\Big(\frac{i - 0.375}{n + 0.25}\Big)$$

$$\begin{aligned} \mathbf{E}X_{(i)} &= \mu + \sigma m_i \\ X_{(i)} &= \mu + \sigma m_i + \epsilon_i \end{aligned}$$

einfaches lineares Regressionsmodell mit Parametern $\mu, \sigma$.
$\mathbf{E}\epsilon_i = 0$, aber die $\epsilon_i$ sind nicht unabhängig.

$$\mathbf{V} := \text{cov}\big(Y_{(i)}, Y_{(j)}\big), \qquad \mathbf{m}' := (m_1, \ldots, m_n)$$

$$\mathbf{X}' := \big(X_{(1)}, \ldots, X_{(n)}\big).$$

Verallgemeinerter Kleinster Quadrat-Schätzer von $\sigma$:

$$\hat{\sigma} = \frac{\mathbf{m}'\mathbf{V}^{-1}\mathbf{X}}{\mathbf{m}'\mathbf{V}^{-1}\mathbf{m}}$$

wird verglichen mit der gewöhnlichen Standardabweichung $s$,

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

# Anpassungstests

Shapiro-Wilk-Statistik

$$W = \frac{\hat{\sigma}^2}{s^2(n-1)} \cdot \frac{(\mathbf{m}'\mathbf{V}^{-1}\mathbf{m})^2}{\mathbf{m}'\mathbf{V}^{-2}\mathbf{m}} = \frac{(\mathbf{h}'\mathbf{X})^2}{\sum_{i=1}^n (X_i - \overline{X})^2 \cdot \mathbf{h}'\mathbf{h}}$$

wobei $\mathbf{h}' = \mathbf{m}'\mathbf{V}^{-1}$ (bekannt, vertafelt).

Wegen $\sum h_i = 0$ folgt:
$W$ ist Quadrat des (empir.) Korrelationskoeffizienten von $\mathbf{h}$ und $\mathbf{X}$:

$$W = \frac{\left(\sum_{i=1}^n (X_i - \overline{X})(h_i - \overline{h})\right)^2}{\sum_{i=1}^n (X_i - \overline{X})^2 \cdot \sum_{i=1}^n (h_i - \overline{h})^2},$$

$$W = \frac{\left(\sum_{i=1}^{n}(X_i - \overline{X})(h_i - \overline{h})\right)^2}{\sum_{i=1}^{n}(X_i - \overline{X})^2 \cdot \sum_{i=1}^{n}(h_i - \overline{h})^2},$$

- Offenbar: $0 \leq W \leq 1$.
- $W \approx 1$ indiziert, dass $\mathbf{h}' = \mathbf{m}'\mathbf{V}^{-1}(\approx 2\mathbf{m}')$:
  ein Vielfaches von $\mathbf{X}$ ist.
  D.h. die Punkte $(m_i, X_{(i)})$ liegen etwa auf einer Geraden, was Normalverteilung indiziert.
- $H_0$ wird ablehnt, falls $W < W_\alpha(n)$.

```
Test_GoF_Shapiro_Wilk.sas
```

# Shapiro-Wilk Test

Scores des 1 Wettkämpfers (5 Preisrichter)

31.2, 31.2, 31.4, 32.0, 33.1
Mit der Prozedur UNIVARIATE erhalten wir $s = 0.80747$ und
mit der Prozedur GPLOT (Option REGEQN) $\hat{\sigma} = 0.805$
im Regressionsmodell $Y_i = \mu + \sigma m_i + \epsilon_i$
Für die Shapiro-Wilk Statistik bekommen wir

$$W = \frac{\hat{\sigma}^2}{s^2} \cdot c = 0.966.$$

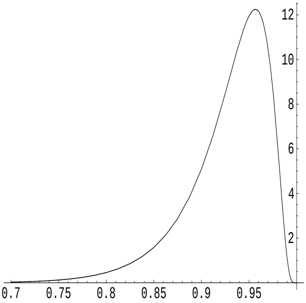
(c: Normierungsfaktor)

# Shapiro-Wilk Test

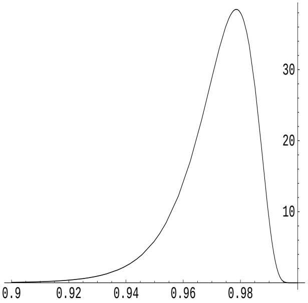
## Approximative Dichtefunktion von $W$ (unter $H_0$)

$$W = \frac{\left(\sum_{i=1}^{n}(X_i - \overline{X})(h_i - \overline{h})\right)^2}{\sum_{i=1}^{n}(X_i - \overline{X})^2 \cdot \sum_{i=1}^{n}(h_i - \overline{h})^2},$$

$n = 10$




$n = 50$



# Anpassungstests

- SAS verwendet eine Approximation von $W$.
- Der Shapiro-Wilk-Test erweist sich für kleinere, mittlere und größere Stichprobenumfänge als geeignetster Test (höchste Güte).
- Früher wurde meist der sogen. $\chi^2$-Anpassungstest verwendet. Dieser hat jedoch geringe Güte.
- $W$ ist etwas besser als A-sq, besser als W-sq, und viel besser als $D$ und $\chi^2$.
- $D$ ist nur für sehr große Stichprobenumfänge zu empfehlen ($n \geq 2000$).

# Anpassungstests

- Man sollte beim Test auf Normalverteilung das Signifikanzniveau auf $\alpha = 0.1$ hochsetzen, insbesondere wenn wenig robuste Tests (die NV verlangen) angewendet werden sollen.

Robuste Tests haben meist geringen Effizienzverlust bei NV.

# Anpassungstests

## Durchführung des Tests auf Normalverteilung

Unter Verwendung von $\hat{\mu}, \hat{\sigma}$:

```
PROC UNIVARIATE NORMAL;
RUN;
```

```
PROC UNIVARIATE;
HISTOGRAM variable / NORMAL;
RUN;
```

mit vorgebenen $\mu, \sigma$:

```
PROC UNIVARIATE;
HISTOGRAM variable / NORMAL(mu=0, sigma=1);
RUN;
```

**Bem.**: Mit der Prozedur UNIVARIATE (Kommando HISTOGRAM) können Sie auch auf andere Verteilungen testen.

# Anpassungstests

## 8.4 Anpassungstests auf weitere Verteilungen

### $\chi^2$-Anpassungstest (Pearson, 1900)

Prinzip: Daten werden in $p$ Klassen eingeteilt.
Klassenhäufigkeiten: $N_i$
theoretische Klassenhäufigkeiten: $np_i$

$$X^2 = \sum_{i=1}^{p} \frac{(N_i - np_i)^2}{np_i}$$

$X^2 \sim \chi^2_{p-1}$ asymptotisch (bei bekannten $\mu, \sigma^2$)
(Fisher, 1922)

$X^2 \sim \chi^2_{p-3}$ approx. (bei 2 zu schätzenden Parametern, ML-Schätzung mit gruppierten Daten oder Minimum-$\chi^2$-Schätzung).

# Anpassungstests

## $\chi^2$-Anpassungstest

Nachteile des $\chi^2$-Anpassungstests

- Wert von $X^2$ abhängig von Klasseneinteilung.
- $\chi^2$- Anpassungstest auf Normalverteilung hat geringe Güte.

Diskrete Verteilungen

Hier kann der $\chi^2$-Anpassungstest genommen werden
(natürliche Klasseneinteilung)
Prozedur FREQ, Option CHISQ

# Anpassungstests

## $\chi^2$-Anpassungstest

Diskrete Gleichverteilung

```
PROC FREQ;
   TABLES var1 /CHISQ;
RUN;
```

Sonstige diskrete Verteilungen

wie oben, zusätzlich sind die Einzelwktn. explizit zu formulieren, /CHISQ TESTP=($p_1, \ldots, p_k$);

`Test_GoF_Poisson.sas`

Anzahlen schon gegeben

Die Variablen, die Anzahlen bezeichnen, werden durch ein WEIGHT-Kommando angegeben.

# Anpassungstests
## EDF-Tests

Stetige Verteilungen

zugelassen sind:

Normal, Gamma, Weibull, Lognormal, Exponential

```
HISTOGRAM var1 / Gamma;
```

`Test_GoF_Darwin_1.sas`