

## Simulationen mit SAS, dargestellt für die Erzeugung der Prüfgröße des David-Tests

Bernd Jäger, Michael Wodny  
Universität Greifswald, Institut für  
Biometrie und Medizinische  
Informatik  
Walther-Rathenau-Str. 48  
17487 Greifswald  
jaeger@biometrie.uni-greifswald.de

Paul Eberhard Rudolph  
Forschungsinstitut für die  
Biologie landwirtschaftlicher  
Nutztiere  
Wilhelm-Stahl-Allee 2  
18196 Dummerstorf  
rudolph@fhn-dummerstorf.de

### Zusammenfassung

Im Statistiksystem SAS stehen viele Grundfunktionen zur Erzeugung von Pseudozufallszahlen zur Verfügung. Mit ihrer Hilfe gelingt es, das Ziehen von Stichproben zu simulieren. Parameterschätzungen und Prüfgrößen statistischer Tests sind Stichprobenfunktionen. Damit ist es möglich, sich Realisierungen einer Schätzung bzw. einer Prüfgröße aus gezogenen Stichproben zu verschaffen. Mit ausreichend großem Simulationsumfang lässt sich auf diese Art eine Häufigkeitsverteilung gewinnen, die der zu beschreibenden Schätz- bzw. Testverteilung nahe kommt.

Die Prüfgrößen des David-Tests für verschiedene Stichprobenumfänge werden auf diese Weise erzeugt, ihre empirischen Quantile ermittelt und mit den Tabellenwerten verglichen.

**Keywords:** David-Test, Simulation mit SAS, Simulation der Testgröße, Vergleich exakte und simulierte Testgröße

## 1 Einleitende Beispiele

(1) Im Technikbereich werden die Parameter der  $N(\mu, \sigma^2)$ -Verteilung aus kleinen Stichproben ( $2 \leq n \leq 10$ ) oftmals nicht durch die erwartungstreuen

Maximum-Likelihood-Schätzung  $m_1 = \left( \sum_{i=1}^n x_i \right) / n$  für  $\mu$  und die Schätzung

$$s_1 = \sqrt{\left( \sum_{i=1}^n (x_i - m_1)^2 \right) / (n-1)}$$
 für  $\sigma$  bestimmt. Die Schätzung  $s_1$  ist erwartungstreu und asymptotisch der Maximum-Likelihood-Schätzung äquivalent. Man bestimmt Mittelwert  $\mu$  und Streuung  $\sigma$  häufig durch die wesentlich leichter zu berechnenden  $m_2 = (x_{(1)} + x_{(n)})/2$  und  $s_2 = (x_{(n)} - x_{(1)})/\alpha_n$ . Dabei bedeuten  $x_{(i)}$  der i-te Wert der der Größe nach geordneten Messreihe und  $\alpha_n$  ein tabellierter Korrekturfaktor. In der Literatur heißen  $m_2$  auch „Spannweitenmitte“ bzw. „midrange“ und  $x_{(n)} - x_{(1)}$  „Spannweite“ bzw. „statistical range“.

(2) Der David-Test, ein Einstichprobentest auf Normalverteilung, hat als Prüfgröße  $D = (X_{(n)} - X_{(1)})/S$  und bemisst damit das Verhältnis von Spannweite und Streuung.

(3) Der Zusammenhang des Schätzproblems 1. und des Testproblems 2. wird durch die Gleichung für  $s_2$  vermittelt. Wird diese nach  $\alpha_n$  umgestellt,  $\alpha_n = (X_{(n)} - X_{(1)})/S$ , erhält man auf der rechten Seite der Gleichung die Prüfgröße des David-Tests. Die in der Literatur mitgeteilten Tabellenwerte sind danach  $\alpha_n = E(D) = E((X_{(n)} - X_{(1)})/S)$ , d.h. der Korrekturfaktor  $\alpha_n$  ist der Erwartungswert der Prüfgröße des David-Tests.

## 2 Theoretische Beschreibung

Aus einer Grundgesamtheit mit stetiger Verteilung wird eine Stichprobe  $(X_1, X_2, \dots, X_n)$  vom Umfang  $n$  gezogen. Diese wird aufsteigend geordnet  $(X_{(1)} \leq X_{(2)} \leq \dots \leq X_{(n)})$ . Wenn die Zufallsgrößen  $X_i$ ,  $(i=1, 2, \dots, n)$ , die Dichtefunktion  $f(x)$  und die Verteilungsfunktion  $F(x)$  besitzen, so sind die Dichtefunktionen der so genannten Ordnungsstatistiken  $X_{(i)}$  gegeben durch

$$f_{X_{(i)}}(x) = \frac{n!}{(i-1)!(n-i)!} F(x)^{i-1} (1-F(x))^{n-i} f(x)$$

und die Verteilungsfunktionen der Ordnungsstatistiken durch

$$F_{X_{(i)}}(x) = \sum_{r=i}^n \binom{n}{r} F(x)^r (1 - F(x))^{n-r}$$

Die Abbildungen 1 und 2 illustrieren für eine Stichprobe vom Umfang  $n = 7$  aus einer  $N(0,1)$ -verteilter Grundgesamtheit Dichte- und Verteilungsfunktionen der Ordnungsstatistiken  $X_{(1)}, X_{(2)}, \dots, X_{(7)}$ . Insbesondere besitzen die Dichtefunktionen der Zufallsgrößen

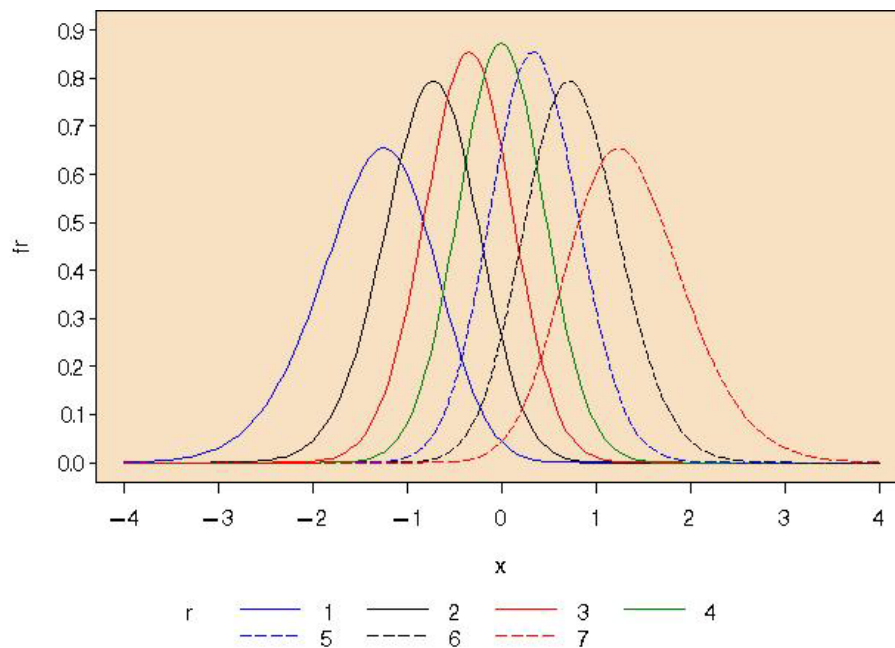
$X_{(1)} = \text{MIN}(X_1, X_2, \dots, X_n)$  und  $X_{(n)} = \text{MAX}(X_1, X_2, \dots, X_n)$ , die für die Spannweite benötigt werden, die Darstellung

$$f_{X_{(1)}}(x) = f_{\text{MIN}}(x) = n \cdot (1 - F(x))^{n-1} f(x) \quad \text{und}$$

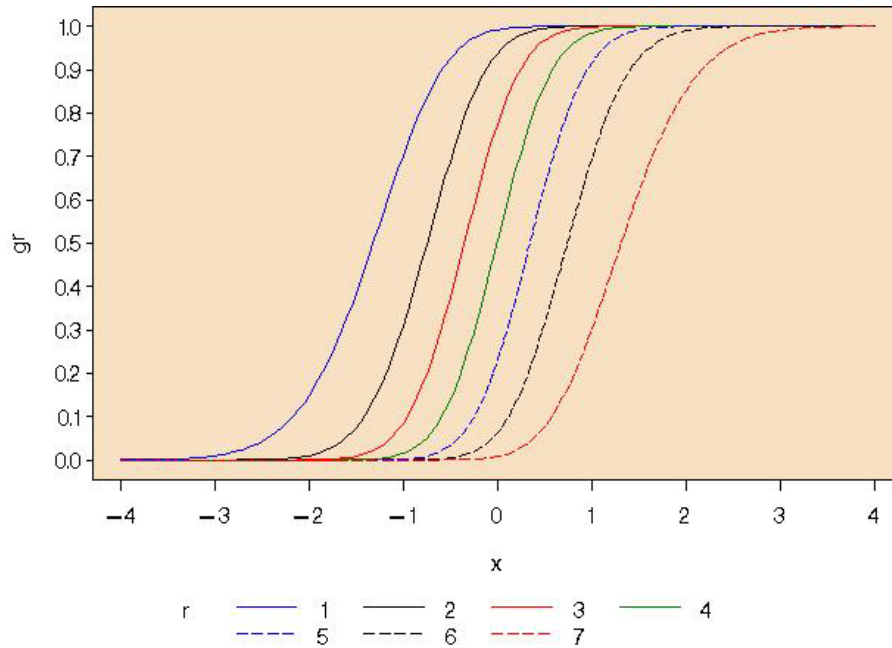
$$f_{X_{(n)}}(x) = f_{\text{MAX}}(x) = n \cdot F(x)^{n-1} f(x).$$

Für die Verteilungsfunktionen gelten

$$F_{X_{(1)}}(x) = F_{\text{MIN}}(x) = 1 - (1 - F(x))^n \quad \text{und} \quad F_{X_{(n)}}(x) = F_{\text{MAX}}(x) = F(x)^n.$$



**Abb. 1:** Dichtefunktionen der Ordnungsstatistiken  $X_{(1)}, X_{(2)}, \dots, X_{(n)}$  für Stichproben des Umfangs  $n = 7$  aus  $N(0,1)$ -verteilter Grundgesamtheit



**Abb. 2:** Verteilungsfunktionen der Ordnungsstatistiken  $X_{(1)}, X_{(2)}, \dots, X_{(n)}$  für Stichproben des Umfangs  $n = 7$  aus  $N(0,1)$ -verteilter Grundgesamtheit

Daraus kann man auf die Dichten und Verteilungen von  $M_2$ ,  $S_2$  und  $D$  schließen, die ihrerseits Funktionen von  $X_{(1)}$  und  $X_{(n)}$  sind. Für die Spannweite

$W = X_{(n)} - X_{(1)}$  erhält man nach Balakrishnan und Cohen (1991) die Dichte

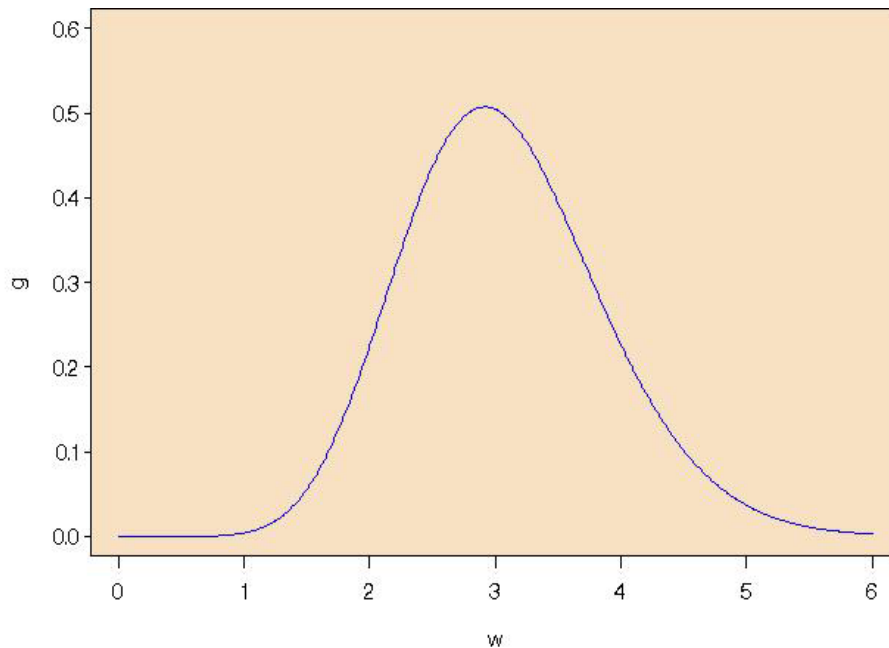
$$g(w) = n(n-1) \int_{-\infty}^{\infty} (F(x+w) - F(x))^{n-2} f(x) f(x+w) dx$$

und die Verteilung

$$G(w) = n \int_{-\infty}^{\infty} (F(x+w) - F(x))^{n-1} f(x) dx$$

in beiden Fällen für  $0 \leq w \leq \infty$ . Diese Formeln erlauben in speziellen Fällen weitere Vereinfachungen. Im betrachteten Beispiel handelte es sich um

die Normalverteilung  $N(0,1)$ , bei der bereits die Verteilungsfunktion  $F(x)$  nur durch numerische Integration zu erhalten ist, und man ist leider auf Näherungsverfahren der Integration angewiesen. Mit dem einfachsten Summationsverfahren (SAS-Programm im Anhang) sind Dichte- und Verteilungsfunktion bestimmt und in den Abbildungen 3 und 4 dargestellt.

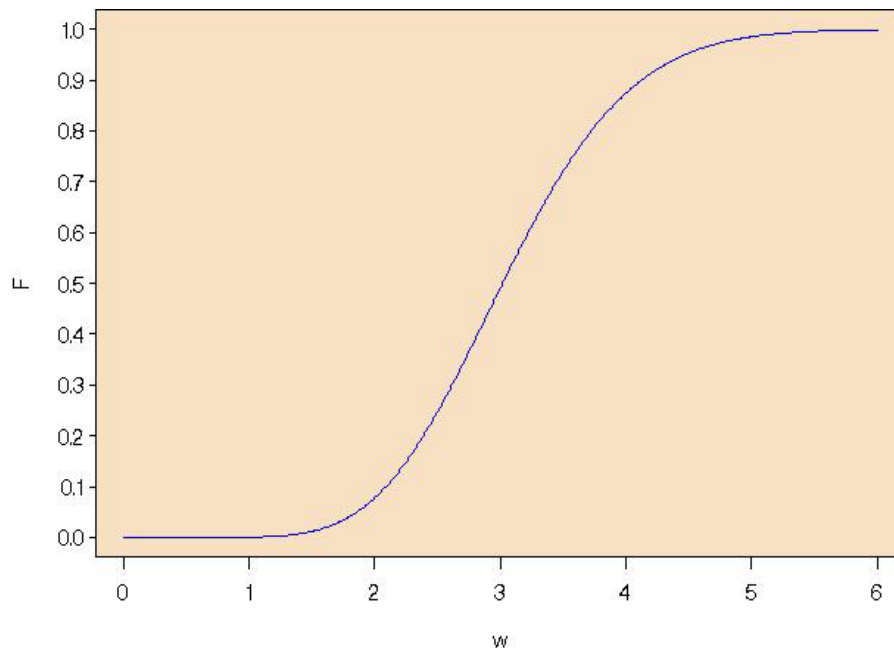


**Abb. 3:** Dichte der Spannweite aus Stichproben vom Umfang  $n = 10$  aus einer Standardnormalverteilung

Darüber hinaus ist für die Berechnung des Erwartungswertes

$$\alpha_n = E(D) = E\left(\frac{X_{(n)} - X_{(1)}}{S}\right) = E(W/S)$$

die Dichte des Quotienten der Zufallsgrößen  $W$  und  $S$  zu bestimmen, eine weitere Schwierigkeit bei der Kalkulation. Auf die vollständige Herleitung der Dichte und Verteilung der David-Prüfgröße soll hier verzichtet werden mit dem Hinweis auf die Originalarbeit von David u. a. (1954).



**Abb. 4:** Verteilung der Spannweite aus Stichproben vom Umfang  $n = 10$  aus einer Standardnormalverteilung

### 3 Das Simulationsexperiment

Das Simulationsexperiment läuft nach folgendem Schema ab:

(1) Gebe die Parameter der Normalverteilung  $\mu$  und  $\sigma$  vor (o.B.d.A. seien  $\mu = 0$  und  $\sigma = 1$ ).

(2) Ziehe eine Stichprobe vom Umfang  $n$  von  $N(0,1)$ -verteilten Zufallszahlen mittels der SAS-Funktion `NORMAL(x)`. Für den Fall  $\mu \neq 0$  und  $\sigma \neq 1$  liefert die Transformation  $\mu + \sigma \cdot \text{NORMAL}(x)$  Zufallszahlen, die  $N(\mu, \sigma^2)$ -verteilt sind.

(3) Berechne daraus Realisierungen der Zufallsgrößen

$$M_1 = \text{MEAN}(X_1, X_2, \dots, X_n), \quad X_{(1)} = \text{MIN}(X_1, X_2, \dots, X_n),$$

$$X_{(n)} = \text{MAX}(X_1, X_2, \dots, X_n), \quad W = X_{(n)} - X_{(1)},$$

$$S = \text{STD}(X_1, X_2, \dots, X_n) \text{ und } D = (X_{(n)} - X_{(1)})/S.$$

(4) Wiederhole die Schritte (1) bis (3) hinreichend oft. Der Simulationsumfang kann so bestimmt werden, dass die Genauigkeitsforderung bezüglich

der Differenz zwischen empirischer Häufigkeitsverteilung und Verteilung mit hoher Wahrscheinlichkeit eingehalten wird. Das liefert der Satz von Kolmogorov und Smirnov. Es wurde ein Simulationsumfang von  $n = 10000$  festgelegt. Die Wahrscheinlichkeit ist 0.05, dass das Supremum der absoluten Differenz der exakten Verteilung und der empirischen Verteilung unterhalb 0.0136 liegt.<sup>37</sup>

Der data step des Simulationsprogramms ist im Anhang 2 zu finden.

## 4 Ergebnisse

Die Schätzung  $M_I$  ist erwartungstreu und hat die Minimalvarianz nach der Rao-Cramer-Ungleichung. Sie ist als Maximum-Likelihood-Schätzung (asymptotisch) normalverteilt.

Da im Simulationsexperiment die Stichprobe vom Umfang  $n = 10$  aus einer Standardnormalverteilung stammte, hat die Zufallsgröße Mittelwert  $M_I$  folglich eine Normalverteilung  $N(\mu, \sigma^2/n) = N(0, 0.1) = N(0, 0.31622^2)$ .

Die erwartete (asymptotische) Normaldichte ist in der Abbildung 5 markiert. Die Zufallsgröße  $M_2$  hat demgegenüber eine deutlich größere Varianz (siehe Abbildung 6).

Die gleichen Minimaleigenschaften sieht man in Abbildung 7 bei den Schätzungen der Standardabweichungen. Maximum-Likelihood-Schätzungen haben asymptotisch stets die Minimalvarianz nach der Rao-Cramer-Ungleichung. Über die Verteilung von  $S_I$  weiß man darüber hinaus, dass  $nS_I^2$  Chi<sup>2</sup>-verteilt mit dem Freiheitsgrad  $n$  ist.<sup>38</sup> Abbildung 9 gibt diese simulierte Chi<sup>2</sup>-Verteilung wieder.

---

<sup>37</sup> Wenn  $S_n(x)$  die empirische Verteilungsfunktion, ermittelt aus einer Stichprobe vom Umfang  $n$ , und  $F(x)$  die unterliegende exakte Verteilung bezeichnet, so gilt für die Prüfgröße

$$D_n = \sup_{-\infty < x < +\infty} |F(x) - S_n(x)|$$

des Kolmogorov-Smirnov-Testes, dass sie wie folgt verteilt ist

$$\lim_{n \rightarrow \infty} Q_n(\lambda) = \lim_{n \rightarrow \infty} P\left(D_n < \frac{\lambda}{\sqrt{n}}\right) = Q(\lambda) = \sum_{k=-\infty}^{\infty} (-1)^k \exp(-2k^2 \lambda^2).$$

<sup>38</sup> Selbstverständlich konvergiert die Schätzung  $S_I$  ebenfalls asymptotisch gegen eine Normalverteilung. Wenn die Stichprobe aus einer  $N(\mu, \sigma^2)$ -Verteilung stammt, ist  $S_I$  nach  $N(\mu, \sigma_{min}^2)$  verteilt, wobei  $\sigma_{min}^2 = 1/(n \cdot I)$  mit  $I$  als so genannter Fisher-Information.

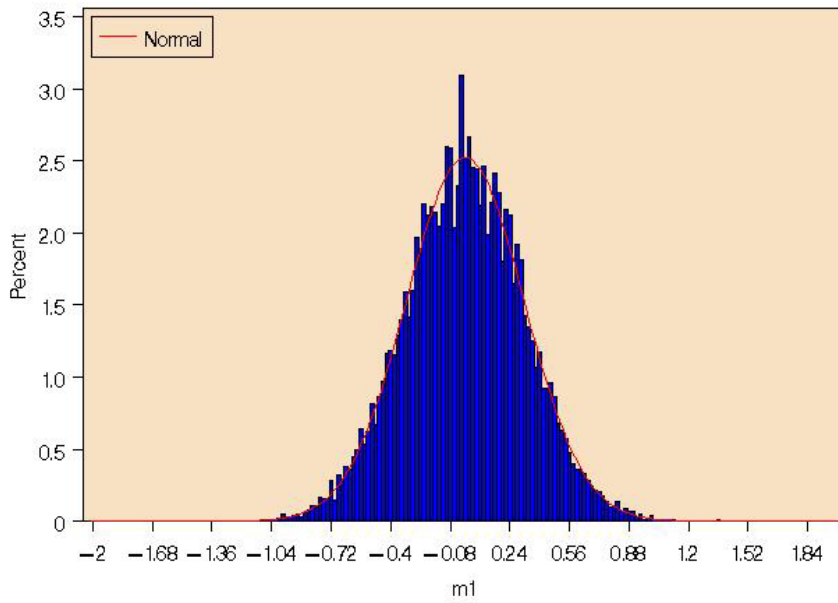


Abb. 5: Simulierte Größen  $M_1$

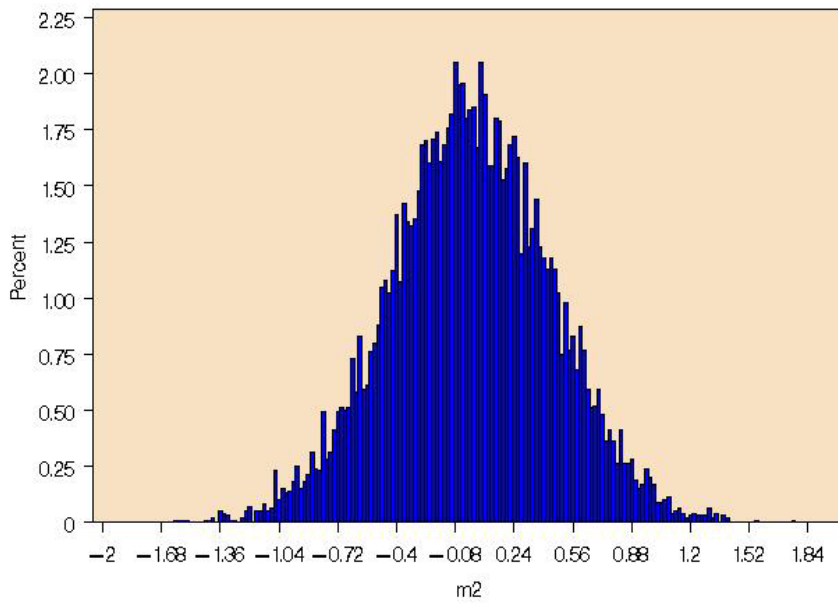


Abb. 6: Simulierte Größen  $M_2$



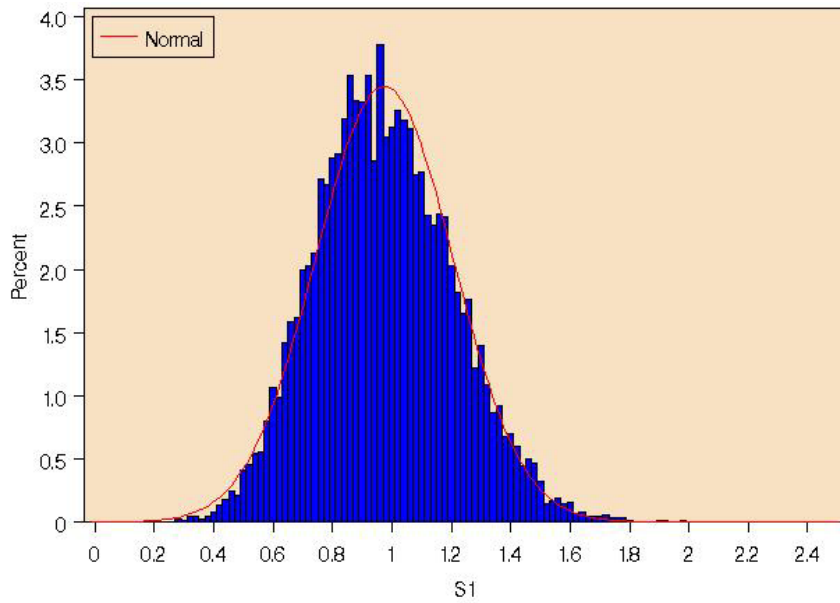


Abb. 7: Simulierte Größe  $S_1$

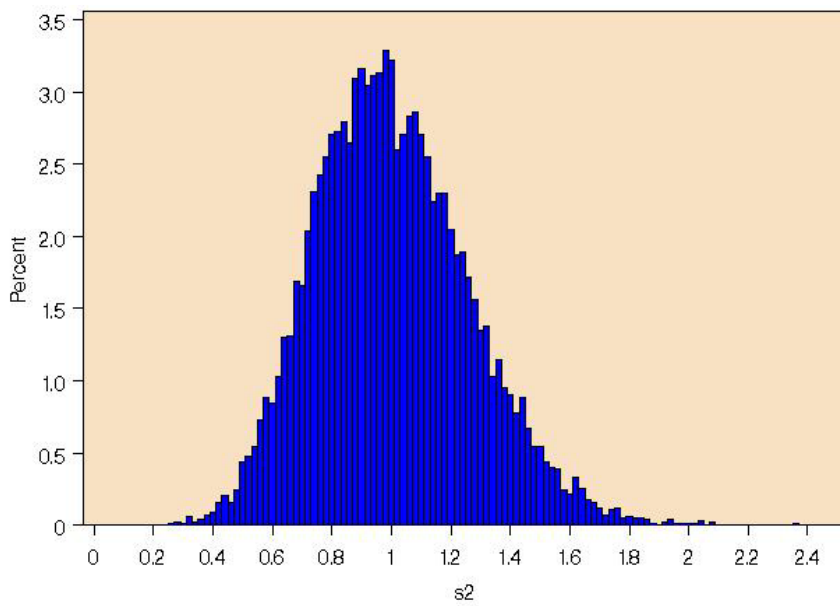
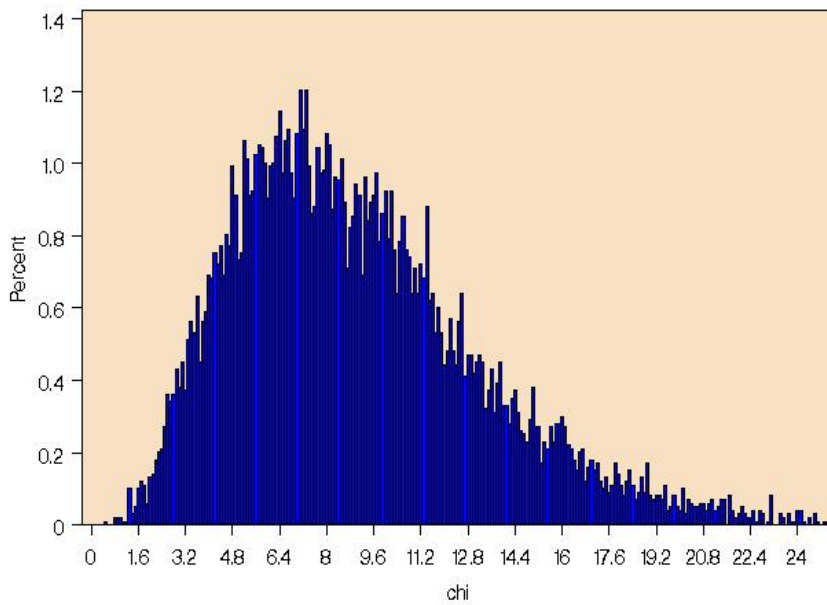
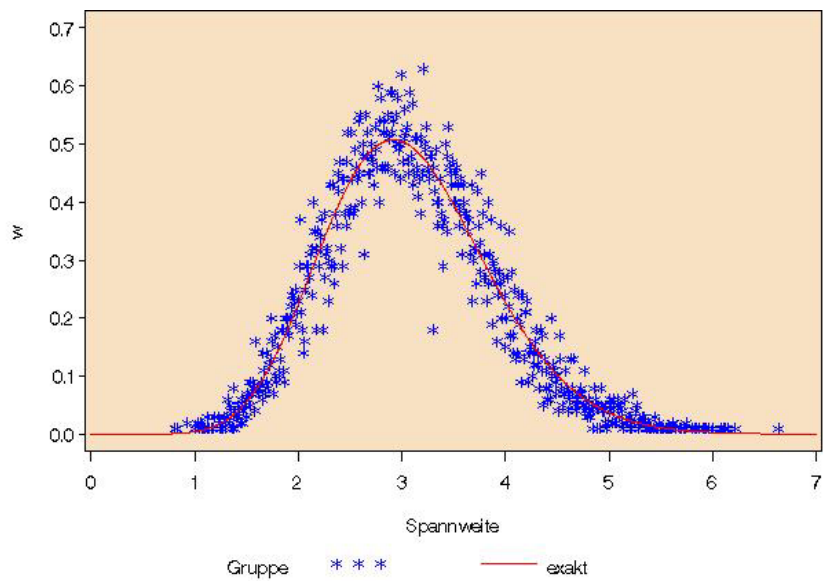


Abb. 8: Simulierte Größe  $S_2 = (X_{(n)} - X_{(1)})/\alpha_n$



**Abb. 9:** Zufallsgröße  $n \cdot S_I^2$  ist Chi<sup>2</sup>-verteilt mit dem Freiheitsgrad n = 10



**Abb. 10:** Simulierte Hf-Verteilung (Sterne) u. exakte Dichte (Linie) der Spannweite bei Stichprobenumfang n = 10 aus N(0,1)-vert. Grundgesamtheit

Die Abbildung 10 stellt die Ergebnisse der Simulation in Bezug auf die Spannweite dar. Man vergleiche diese Häufigkeitsverteilung mit der zugehörigen exakten Dichte der Spannweite in Abbildung 3.

Mit der PROC UNIVARIATE lassen sich die Quantile der simulierten David-Testgröße ermitteln. Diese stimmen sehr gut mit den in der Literatur mitgeteilten exakt berechneten Quantilen überein. Programmtechnisch wurden das Simulationsprogramm und die PROC UNIVARIATE in ein Macro eingebaut. Eingabevariable ist der Stichprobenumfang und Ausgabevariable der Vektor der Quantile.

Tabelle 1 gibt in Abhängigkeit vom Stichprobenumfang  $n$  die empirischen  $Q_{0.005}$ ,  $Q_{0.01}$ ,  $Q_{0.025}$ ,  $Q_{0.05}$ ,  $Q_{0.10}$ ,  $Q_{0.90}$ ,  $Q_{0.95}$ ,  $Q_{0.975}$ ,  $Q_{0.99}$  und  $Q_{0.995}$  an. Die exakten Quantile aus der Literatur sind dabei den simulierten gegenübergestellt. Abweichungen treten in wenigen Fällen an der zweiten Dezimalstelle auf, die Regel sind Abweichungen an der dritten Dezimalstelle.

Die Vorgehensweise bei der Berechnung der Tabellenwerte nach David u.a. (1954) wurde referiert von M.P. Geppert im Zentralblatt für Mathematik, vol.56, p.366 wie folgt: „Zu dem Zwecke werden aus den bekannten Momenten von  $W$  (Spannweite) und  $S$  (Streuung) diejenigen von  $D$  exakt berechnet und Pearson-Kurven, deren erste vier Momente mit denjenigen von  $D$  übereinstimmen, zur Approximation der unbekannt exakten Verteilung von  $D$  benutzt.“

Bei Dichten aus dem Pearson-System besteht eine eindeutige Beziehung zwischen den Koeffizienten der beschreibenden Differentialgleichung und den ersten vier Momenten. Ab fünften Moment aufwärts sind die Verteilungen von  $D$  und ihrer Pearson-Approximation in der Regel different. Deshalb ist die Bezeichnung „exakte Quantile“ in der Literatur vermessen. Die exakten Quantile sind nicht die von  $D$ , sondern die der zugehörigen Pearson-Verteilung. Bedenkt man darüber hinaus, dass der Fehler zwischen der simulierten und der exakten David-Verteilung mittels des Satzes von Kolmogorov und Smirnov quantifiziert und durch Vergrößern des Simulationsumfangs beliebig verkleinert werden kann, sollte man die Simulationsmethode favorisieren, weil im Gegensatz zur klassischen Vorgehensweise ihre Approximation kalkulierbar ist.

**Tabelle 1:** Exakte und simulierte Quantile des David-Tests für Stichprobenumfänge  $n = 3$  bis 20, 25, 30, 35, 40 und 50

Stichprobenumfang	Quantile Q									
	0.005	0.01	0.025	0.05	0.10	0.90	0.95	0.975	0.99	0.995
3	1.735	1.737	1.745	1.758	1.782	1.997	1.999	2.000	2.000	2.000
simuliert	1.734	1.737	1.746	1.759	1.785	1.997	1.999	2.000	2.000	2.000
4	1.83	1.87	1.93	1.98	2.04	2.409	2.429	2.439	2.445	2.447
simuliert	1.808	1.846	1.910	1.974	2.049	2.408	2.429	2.440	2.446	2.448
5	1.98	2.02	2.09	2.15	2.22	2.712	2.753	2.782	2.803	2.813
simuliert	1.997	2.032	2.082	2.142	2.221	2.709	2.754	2.781	2.801	2.811
6	2.11	2.15	2.22	2.28	2.37	2.949	3.012	3.056	3.095	3.115
simuliert	2.117	2.150	2.216	2.275	2.361	2.947	3.014	3.055	3.097	3.115
7	2.22	2.26	2.33	2.40	2.49	3.143	3.222	3.282	3.338	3.369
simuliert	2.215	2.263	2.340	2.406	2.483	3.142	3.221	3.284	3.339	3.368
8	2.31	2.35	2.43	2.50	2.59	3.308	3.399	3.471	3.543	3.585
simuliert	2.316	2.356	2.432	2.504	2.594	3.315	3.401	3.472	3.542	3.582
9	2.39	2.44	2.51	2.59	2.68	3.449	3.552	3.634	3.720	3.772
simuliert	2.392	2.438	2.515	2.590	2.684	3.447	3.553	3.630	3.713	3.763
10	2.46	2.51	2.59	2.67	2.76	3.57	3.685	3.777	3.875	3.935
simuliert	2.452	2.502	2.597	2.672	2.768	3.583	3.693	3.783	3.882	3.930
11	2.53	2.58	2.66	2.74	2.84	3.68	3.80	3.903	4.012	4.079
simuliert	2.518	2.574	2.659	2.742	2.834	3.676	3.797	3.893	3.998	4.064
12	2.59	2.64	2.72	2.80	2.90	3.78	3.91	4.02	4.134	4.208
simuliert	2.574	2.629	2.719	2.803	2.908	3.782	3.902	4.012	4.119	4.191
13	2.64	2.70	2.78	2.86	2.96	3.87	4.00	4.12	4.244	4.325
simuliert	2.630	2.689	2.773	2.858	2.954	3.874	3.998	4.102	4.236	4.321
14	2.70	2.75	2.83	2.92	3.02	3.95	4.09	4.21	4.34	4.431
simuliert	2.696	2.739	2.825	2.906	3.021	3.944	4.088	4.202	4.333	4.418

Stich- proben- umfang	Quantile Q									
	0.005	0.01	0.025	0.05	0.10	0.90	0.95	0.975	0.99	0.995
15	2.74	2.80	2.88	2.97	3.07	4.02	4.17	4.29	4.44	4.53
simuliert	2.739	2.790	2.875	2.964	3.068	4.020	4.169	4.300	4.441	4.535
16	2.79	2.84	2.93	3.01	3.12	4.09	4.24	4.37	4.52	4.62
simuliert	2.768	2.845	2.931	3.022	3.122	4.091	4.242	4.377	4.526	4.632
17	2.83	2.88	2.97	3.06	3.17	4.15	4.31	4.44	4.60	4.70
simuliert	2.824	2.875	2.971	3.060	3.165	4.152	4.315	4.451	4.622	4.751
18	2.87	2.92	3.01	3.10	3.21	4.21	4.37	4.51	4.67	4.78
simuliert	2.859	2.910	3.003	3.099	3.206	4.219	4.376	4.506	4.662	4.769
19	2.90	2.96	3.05	3.14	3.25	4.27	4.43	4.57	4.74	4.85
simuliert	2.903	2.958	3.054	3.142	3.249	4.277	4.449	4.599	4.758	4.874
20	2.94	2.99	3.09	3.18	3.29	4.32	4.49	4.63	4.80	4.91
simuliert	2.915	2.992	3.082	3.171	3.286	4.321	4.489	4.648	4.837	4.950
25	3.09	3.15	3.24	3.34	3.45	4.53	4.71	4.87	5.06	5.19
simuliert	3.089	3.153	3.2431	3.342	3.459	4.540	4.708	4.861	5.051	5.196
30	3.21	3.27	3.37	3.47	3.59	4.70	4.89	5.06	5.26	5.40
simuliert	3.200	3.274	3.382	3.477	3.591	4.691	4.884	5.045	5.231	5.375
35	3.32	3.38	3.48	3.58	3.70	4.84	5.04	5.21	5.42	5.57
simuliert	3.310	3.374	3.486	3.579	3.705	4.829	5.014	5.182	5.381	5.545
40	3.41	3.47	3.57	3.67	3.79	4.96	5.16	5.34	5.56	5.71
simuliert	3.390	3.466	3.571	3.667	3.793	4.966	5.174	5.366	5.592	5.740
50	3.56	3.62	3.73	3.83	3.95	5.14	5.35	5.54	5.77	5.93
simuliert	3.566	3.635	3.751	3.845	3.973	5.136	5.363	5.554	5.754	5.933

## Literatur

- [1] Balakrishnan, N., Cohen, A. C. (1991). Order Statistics and inference, Academic Press, Inc., San Diego
- [2] David, H. A., Hartley, H. O., Pearson, E. S. (1954). The distribution of the ratio in a single normal sample of range to standard deviation. Biometrika 41, 482-493
- [3] Efron, B., Tibshirani, R.J. (1993). An introduction to the bootstrap, Chapman & Hall, Inc., New York, London,
- [4] Fisz, M.: Wahrscheinlichkeitsrechnung und mathematische Statistik, Dt. Verl. d. Wiss., Berlin, 1989
- [5] Pearson, E. S., Stephens, M. A. (1964). The ratio of range to standard deviation in the same normal sample. Biometrika 51, 484-487
- [6] SAS Institute Inc. (1989), SAS/STAT User's Guide, Version 6, Fourth Edition, Volume 1, Cary, NC: SAS Institute Inc.
- [7] <http://www.medizin.uni-greifswald.de/biometrie/statist/>

## Anhang

### Anhang 1: Programm zur Berechnung von Dichte und Verteilung der Ordnungsstatistiken

```
/* Dichtefunktionen und Verteilungsfunktionen
   der Ordnungsstatistiken von Normalverteilungen
   Balakrishnan, N. und Cohen, A.C.:
   Order statistics and inference, Academic Press, Inc.
   Boston (1990) */

data balakrishnan;
n=7; /* Stichprobenumfang */
Pi=3.1415927;
do x=-4 to 4 by 0.05; /* Definitionsbereich */
do r=1 to n;
f_r=gamma(n+1)/(Gamma(r)*Gamma(n-r+1))
    *Probnorm(x)**(r-1)*(1-Probnorm(x))**(n-r)
```

```

*1/SQRT(2*Pi)*EXP(-x**2/2); /* Dichten */
gr=0;
do rr=r to n;
gr=gr+Gamma(n+1)/(Gamma(rr+1)*Gamma(n-rr+1))
*Probnorm(x)**rr*(1-Probnorm(x)**(n-rr));
end; /* Verteilungsfunktionen */
output;
end;
end;
run;

```

## Anhang 2: data-step für die Simulation

```

data simul;
my=0;
sigma=1;
n=10;
do i=1 to 10000;
x1=my+sigma*NORMAL(-1);
x2=my+sigma*NORMAL(-1);
x3=my+sigma*NORMAL(-1);
x4=my+sigma*NORMAL(-1);
x5=my+sigma*NORMAL(-1);
x6=my+sigma*NORMAL(-1);
x7=my+sigma*NORMAL(-1);
x8=my+sigma*NORMAL(-1);
x9=my+sigma*NORMAL(-1);
x10=my+sigma*NORMAL(-1);
Ma=MAX(x1,x2,x3,x4,x5,x6,x7,x8,x9,x10); /* Maximum */
Mi=MIN(x1,x2,x3,x4,x5,x6,x7,x8,x9,x10); /* Minimum */
m1=MEAN(x1,x2,x3,x4,x5,x6,x7,x8,x9,x10);
/* MLH-Schätzung für my */
m2=(Ma+Mi)/2; /* alternative Schätzung für my */
S1=STD(x1,x2,x3,x4,x5,x6,x7,x8,x9,x10);
/* s1 MLH-Schätzung für sigma, */
s2=(Ma-Mi)/3.077; /* alternative Schätzung für sigma */
chi=n*s1**2; /* n*s1**2 ist CHI2-verteilt, df=n */
d=(Ma-Mi)/s1; /* Prüfgröße des David-Tests */
Spannweite=Ma-Mi; /* Schätzung der Spannweite */
output;
end;
run;

```

### Anhang 3: Berechnung von Dichte und Verteilung der Spannweite

```
/* Berechnung der Dichte- und der Verteilungsfunktion
   der Stichprobenspannweite aus N(0,1)-Verteilung */
data range;
keep w g f;
pi=3.1415927;
n=10; /* Stichprobenumfang */
F=0;
sw1=0.005; /*Schrittweite bei der Spannweite */
sw2=0.005; /*Schrittweite bei numer.Integration*/
do w=0 to 12 by sw1;
/* Spannweite wird als von 0 bis 10
   variierend angenommen */
integral=0;
do x=-20 to 20 by sw2; /* Stützstellen Integration*/
y1=1/SQRT(2*pi)*EXP(-x**2/2);
y2=1/SQRT(2*pi)*EXP(-(x+w)**2/2);
y3=(PROBNORM(x+w)-PROBNORM(x))**(n-2);
z=y1*y2*y3;
integral=integral+z*sw2;
/* Berechnung des Integrals durch Streifensummation */
end;
g=n*(n-1)*integral;
F=F+g*sw1;
output;end;
run;
```