

Unimodalität und Unimodalitätstests

Gabler, Siegfried; Borg, Ingwer

Veröffentlichungsversion / Published Version

Zeitschriftenartikel / journal article

Zur Verfügung gestellt in Kooperation mit / provided in cooperation with:

GESIS - Leibniz-Institut für Sozialwissenschaften

Empfohlene Zitierung / Suggested Citation:

Gabler, S., & Borg, I. (1996). Unimodalität und Unimodalitätstests. *ZUMA Nachrichten*, 20(38), 33-44. <https://nbn-resolving.org/urn:nbn:de:0168-ssoar-208834>

Nutzungsbedingungen:

Dieser Text wird unter einer Deposit-Lizenz (Keine Weiterverbreitung - keine Bearbeitung) zur Verfügung gestellt. Gewährt wird ein nicht exklusives, nicht übertragbares, persönliches und beschränktes Recht auf Nutzung dieses Dokuments. Dieses Dokument ist ausschließlich für den persönlichen, nicht-kommerziellen Gebrauch bestimmt. Auf sämtlichen Kopien dieses Dokuments müssen alle Urheberrechtshinweise und sonstigen Hinweise auf gesetzlichen Schutz beibehalten werden. Sie dürfen dieses Dokument nicht in irgendeiner Weise abändern, noch dürfen Sie dieses Dokument für öffentliche oder kommerzielle Zwecke vervielfältigen, öffentlich ausstellen, aufführen, vertreiben oder anderweitig nutzen.

Mit der Verwendung dieses Dokuments erkennen Sie die Nutzungsbedingungen an.

Terms of use:

This document is made available under Deposit Licence (No Redistribution - no modifications). We grant a non-exclusive, non-transferable, individual and limited right to using this document. This document is solely intended for your personal, non-commercial use. All of the copies of this documents must retain all copyright information and other information regarding legal protection. You are not allowed to alter this document in any way, to copy it for public or commercial purposes, to exhibit the document in public, to perform, distribute or otherwise use the document in public.

By using this particular document, you accept the above-stated conditions of use.

UNIMODALITÄT UND UNIMODALITÄTSTESTS

SIEGFRIED GABLER UND INGWER BORG

Kennzahlen spielen in der Beschreibung von Daten eine wichtige Rolle, wobei der Mittelwert als Lageparameter die beliebteste Kennzahl ist. Die Aussagekraft des arithmetischen Mittels sinkt allerdings gravierend, wenn die Verteilung etwa U-förmig oder allgemeiner, mehrgipfelig ist. Klassierte Daten lassen sich auch in Histogrammen darstellen. Ob aus ihrer Darstellung aber auf die der Population zugrunde liegende Verteilung geschlossen werden kann, hängt häufig von der gewählten Intervalleinteilung ab. Keines der gängigen Statistik-Lehrbücher beschreibt einen für die Frage nach der Unimodalität angemessenen Test. Überraschend ist auch, daß keines der großen Statistikpakete einen solchen Test anbietet. Alles, was man findet, sind Tests, bei denen die Anpassung an eine fest vorgegebene Verteilung vorgenommen wird. Um die grundsätzliche Wichtigkeit der Thematik in das Blickfeld der empirischen Sozialwissenschaften zu rücken, beschreiben wir einen wenig bekannten Test auf Unimodalität, den sogenannten DIP-Test. Anhand von Skalenwerten erläutern wir das Vorgehen.

Index statistics play an important role in describing data. The most important such statistic in practice is the mean. It is often overlooked, however, that the interpretability of the mean depends strongly on the shape of the distribution. If the distribution is multi-peaked or U-shaped, the mean is misleading. To check the shape of the distribution by using histogram plots is tricky, because histograms can change substantially if the number of bars is slightly modified. None of the major statistics textbooks describes an inferential test for unimodality, and none of the major statistics program packages provides such a test either. What they offer are tests for particular distributions such as the normal. To direct the attention of social scientists to the importance of the unimodality issue, we here describe a little known unimodality test, the DIP test. An application is given, and practical issues are discussed.

1. Zur Bedeutsamkeit von Maßen zentraler Tendenz

Die vermutlich meistverwendeten Statistiken in der sozialwissenschaftlichen Berichterstattung sind Mittelwerte wie etwa die bekannten durchschnittlichen Popularitätswerte

für Politiker. Eine sinnvolle Interpretation des Mittelwertes unterstellt, daß dieser tatsächlich eine zentrale Tendenz beschreibt in dem Sinn, daß die Beobachtungen um diesen Wert herum klumpen. Wenn sich dagegen z.B. die Präferenzen für politische Entscheidungen auf einer Links-Rechts-Dimension polarisieren, d.h. die Verteilung U-förmig ist, dann ist der Mittelwert irreführend, weil nur ganz wenige oder überhaupt keine der Befragten die Parteien der Mitte bevorzugen, wie der Mittelwert nahelegen würde. Es ist in diesem Fall offensichtlich, daß aus der naiven Interpretation des Mittelwertes als dem mehr oder weniger "typischen" Wert grobe Fehler in der Vorhersage des Wahlscheids entstehen können.

Vielverwendet werden in den Sozialwissenschaften auch kumulierte Prozentangaben, um die Verteilung auf mehrstufigen Skalen zu charakterisieren. Berichtet wird dann etwa, daß sich 60% "zustimmend" und 40% "ablehnend" zu einem bipolaren Item geäußert haben. Auch bei diesen Statistiken tritt das Problem der zentralen Tendenz auf, weil Zustimmungs- und Ablehnungsprozente natürlicherweise als Schnittstellen einer kontinuierlichen, eingipfeligen Verteilung gedeutet werden (Thurstone 1927). Das ist im allgemeinen auch eine empirisch recht gut begründete Annahme, weil sich für Einstellungs- und Meinungsitems zeigt, daß Zustimmungs- und Ablehnungsprozente hoch mit Mittelwerten korrelieren. Wenn also z.B. 60% einem Item zustimmen und 40% dieses ablehnen, dann kann man hieraus schließen, daß die Schnittstelle zwischen Zustimmung und Ablehnung bei etwa $z=+0.26$ liegt (für eine z-standardisierte Variable). Insofern kann man derartige Prozente auch an Stelle von Mittelwerten in der Sozialberichterstattung verwenden (Borg 1989).

Die Voraussetzung für eine sinnvolle Anwendung solcher Maße ist nicht unbedingt eine Normalverteilung, sondern vor allem die Frage, ob man davon ausgehen kann, daß die beobachtete Verteilung aus einer eingipfeligen Population stammt. Ist das nicht der Fall, ist es in jedem Fall überlegenswert, ob es nicht theoretisch fruchtbarer ist, die Population in verschiedene Klassen zu partitionieren. Coombs/Avrunin (1977) haben beispielsweise eine Theorie für Präferenzverhalten entwickelt, die die Eingipfeligkeit von Präferenzfunktionen über einem Merkmalskontinuum aus zwei zugrundeliegenden Funktionen über diesem Merkmalskontinuum ableitet. Beobachtet man mehrgipfelige Präferenzverteilungen, dann sind nach dieser Theorie die Objekte des Auswahlraums grundsätzlich nicht vergleichbar, wie z.B. die Objekte "Eistee" und "heißer Tee", die eine zweigipfelige Verteilung über der Merkmalsdimension "Temperatur" erzeugen. In diesem Fall sind die Daten auch nicht sinnvoll in einem Unfolding-Modell darstellbar (Coombs 1964).

Keines der gängigen Statistik-Lehrbücher beschreibt einen für die Frage nach der Unimodalität der einer beobachteten Verteilung zugrundeliegenden Verteilung angemessenen Test. Noch wichtiger ist, daß auch keines der großen Statistikpakete (z.B. SPSS, SYSTAT, STATISTICA, SAS) einen solchen Test anbietet. Beschrieben bzw. angeboten werden nur besondere Tests, wie z.B. solche zur Prüfung auf Normalität der Verteilung. In der Praxis werden sie aber selten angewendet, wohl deshalb, weil eine zumindest approximative Normalität der Verteilung in den meisten Fällen einfach unterstellt wird. Diese Unterstellung ist sicher nicht unberechtigt. Sie kann aber im Einzelfall falsch sein und führt dann zu schwerwiegenden Fehlinterpretationen der Statistiken. Die folgende Darstellung soll deshalb auch dazu dienen, die grundsätzliche Wichtigkeit der Thematik in das Blickfeld der empirischen Sozialwissenschaften zu rücken.

2. Unimodalität bei Histogrammen

Eine beliebte grafische Darstellung von Daten sind die Histogramme. Sie sind als universelle grafische Darstellungsform für Verteilungen in der Statistik in praktisch allen Statistikprogrammen zu finden. In der Regel werden auf der horizontalen Achse Intervalle gleicher Breite verwendet und die Höhe der Balken durch die absolute oder relative Zahl der Beobachtungen im entsprechenden Intervall ermittelt. Gibt es nur wenige Ausprägungen, wie bei den Likertskalen, die in den Sozialwissenschaften weit verbreitet sind, geben die Histogramme ein genaues Abbild der Häufigkeiten wider, da eine Gruppierung der Daten nicht notwendig ist. Streuen die Daten aber innerhalb eines Intervalls, gehen wertvolle Informationen bei der Darstellung der Daten durch Histogramme verloren.

Unimodalität eines Histogramms ist dann gegeben, wenn die Höhen der Balken von links nach rechts erst bis zu einem bestimmten Punkt wachsen und dann wieder sinken. In diesem Sinne liefern die Daten des ALLBUS für viele Einstellungstems unimodale Histogramme. Da viele der bekannten diskreten Verteilungen (Binomial, Poisson) bei dieser Definition unimodal sind, kann man natürlich auch durch einen Anpassungstest überprüfen, ob die Daten aus einer solchen Verteilung stammen. Vielfach hat man aber die Vorstellung, daß in der Population eine stetige Verteilung vorliegt und die beobachteten Daten Realisierungen aus dieser Verteilung sind. Bekannte Anpassungstests prüfen dann etwa auf Normalität. Stellt man die Daten in Form von Histogrammen dar, geben sie uns nur noch eine mehr oder weniger gute Annäherung für die in der Population zugrunde liegende Verteilung. Ob Unimodalität des Histogramms zum Vorschein kommt oder nicht, ist dabei häufig nur eine Frage der Anzahl der gewählten Intervalle.

3. Eine illustrative Anwendung

Betrachten wir eine Anwendung. Borg/Braun (1992) berichten die Skalenwerte von 125 Mitarbeitern eines Gießener Textilbetriebes auf einer Skala zur Messung der subjektiven Unsicherheit der Arbeitsstelle (SUSA) (Borg 1992). Die Skala unterscheidet zwei Faktoren: 'Bedenken' und 'Sorgen'. Bedenken werden erfaßt mit sechs Items über den wahrscheinlichen Erhalt der Arbeitsstelle, der Gewißheit über die eigene Karriere, der Zuversicht, neue Technologien meistern zu können, usw. Sorgen werden gemessen mit drei Items, bei denen es um die psychische Belastung geht, die der mögliche Verlust der Arbeitsstelle bewirkt. Empirisch zeigt sich bei den 103 non-missing data für 'Sorgen' und 105 non-missing für 'Bedenken', daß Sorgen nur dann auftreten, wenn zumindest gewisse Bedenken gegeben sind. Es gibt aber Personen, die große Bedenken haben, sich aber trotzdem keine Sorgen machen.

Für die Stichprobe ergeben sich die in Abbildung 1 gezeigten Verteilungen der Skalenwerte. Die Abbildungen zeigen Histogramme mit angepaßten Normalverteilungen (fette Glockenkurven) und jeweils einer zweiten gefitteten Funktion (gestrichelte Linien), die eine bessere Anpassung erlaubt (Kernel-Funktionen). Beide Abbildungen wurden mit SYSTAT erzeugt, wobei der Streubereich der X-Achse und der maximale Y-Wert vorgegeben wurden, um bessere Vergleichbarkeit zu erreichen. Die Intervallbildung auf der X-Achse wurde dagegen dem Programm überlassen. Wie man sieht, sind beide Verteilungen nicht sehr normal, vor allem nicht die von Sorgen. Bei Sorgen sieht die Verteilung eher tendenziell zwei-modal aus, wie auch die Kernel-Funktion unterstreicht. Der Kolmogorov-Smirnov-Anpassungstest, kurz K-S Test, lehnt die Normalverteilungshypothese bei Sorgen mit $p < .05$ ab, bei Bedenken aber nicht ($p < .15$). Die letzte Aussage ist allerdings etwas irreführend, weil beim K-S Test der Erwartungswert und die Varianz der Normalverteilung festgelegt werden müssen. Daß bei Bedenken Unimodalität nicht abgelehnt wird, liegt nur daran, daß die Parameter der Verteilung durch die entsprechenden Stichprobengrößen vom Programm automatisch festgelegt wurden. Werden die Daten erst standardisiert und auf Standardnormalverteilung getestet, erhält man den entsprechenden Test mit Lilliefors-Schranken. Für ihn gilt in beiden Fällen $p < .01$. In SYSTAT werden der K-S Test und der Lilliefors Test angeboten und es liegt am Nutzer, den richtigen zu wählen.

Wählt man statt der von SYSTAT intern bestimmten 15 Intervalle für die X-Achse nur 14 solcher Intervalle, also ein geringfügig gröberes Raster, dann ergibt sich Abbildung 2.

Abbildung 1: Häufigkeitsverteilungen (Histogramme über 15 Intervallen der X-Achse) der Skalenwerte 'Bedenken' und 'Sorgen', zusammen mit angepassten Normalverteilungen und Kernel-Funktionen

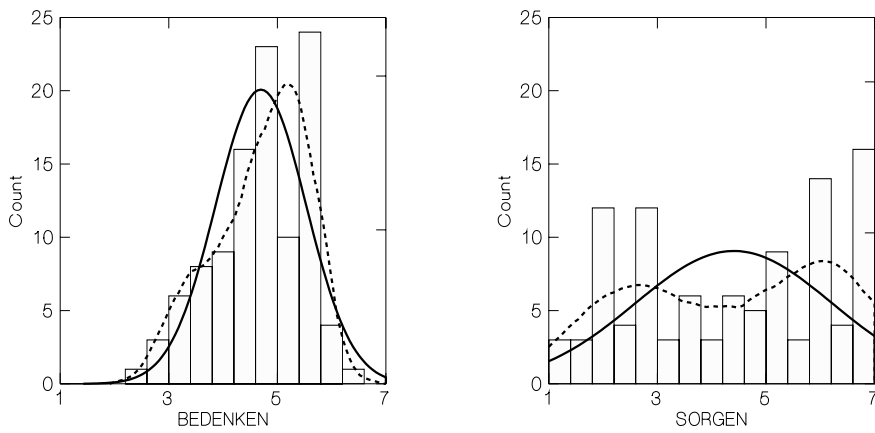
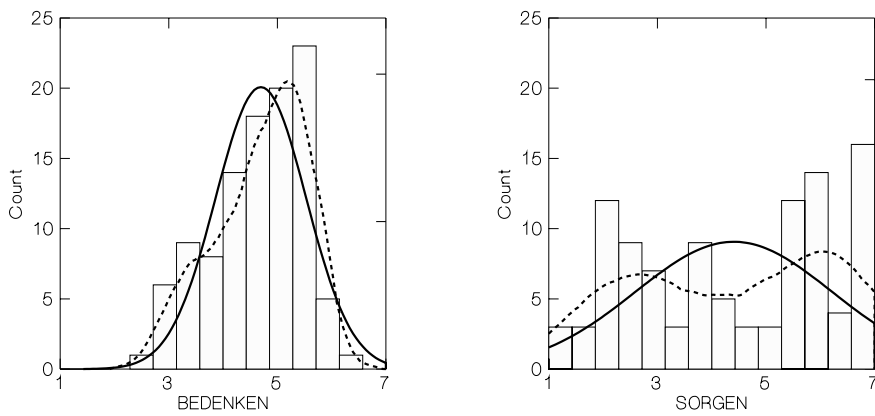


Abbildung 2: Häufigkeitsverteilungen wie in Abbildung 1, mit 14 anstatt 15 Intervallen auf der X-Achse



Man sieht, daß diese kleine Änderung zu einem deutlich anderen Bild der Verteilung der Skalenwerte für Bedenken führt: Der tiefe Einschnitt zwischen den beiden höchsten Säulen in Abbildung 1 ist nun völlig verschwunden ist. (Die gefitteten Funktionen sind dagegen exakt gleich, weil sie nicht dem Histogramm, sondern den Rohdaten angepaßt werden.) Die Programmpakete SPSS und STATISTICA zeigen (per Voreinstellung) ein noch gröberes Raster als 14 Intervalle. Dadurch wird der Eindruck noch weiter verstärkt, daß Bedenken weitgehend normal verteilt sind, während man in Abbildung 1 doch immerhin sieht, daß man auch hier Zweifel haben könnte, ob die Population normal verteilt ist. Insgesamt erkennt man durch dieses Beispiel das grundsätzliche Problem, daß eine Überprüfung der Normalität oder der Eingipfeligkeit mit Histogrammen bei Verwendung verschiedener Auswertungsprogramme zu verschiedenen Entscheidungen führen kann. Das Problem der Rasterung tritt in jedem Fall auf. Im folgenden wird ein Test zur Prüfung der Unimodalität von Verteilungen beschrieben und illustriert.

4. Unimodalität von Verteilungen

Es gibt nur wenige Tests, die die Existenz von mehreren Gipfeln in einer Verteilung überprüfen und nicht jeder dieser Tests ist hinreichend verläßlich unter Bedingungen, die man in der Praxis nicht von vornherein ausschließen kann (Hartigan/Hartigan 1985a). Robust und gleichzeitig transparent ist dagegen der DIP-Test von Hartigan/Hartigan (1985a) für eindimensionale Verteilungen¹⁾. Er basiert auf der DIP-Statistik als Prüfgröße, die die Multimodalität einer Stichprobe mißt als maximale Differenz aller Punkte der empirischen Verteilungsfunktion von der unimodalen Verteilungsfunktion, die diese maximale Differenz minimiert. Die Autoren vermuten, daß die Gleichverteilung die unimodale Verteilung ist, die am schlechtesten von den multimodalen Verteilungen unterschieden werden kann. Sie wird daher als Nullverteilung in der Nullhypothese verwendet.

Es sei F eine Verteilungsfunktion. Dann heißt F *unimodal* mit Modus m , wenn eine Zahl m existiert derart, daß F im Intervall $(-\infty, m]$ konvex und im Intervall $[m, \infty)$ konkav ist. Der Wert m ist nicht notwendig eindeutig. Eine unimodale stetige Verteilung hat also eine Dichte, die rechts vom Modalwert monoton fallend und links vom Modalwert monoton steigend ist. Stetige Verteilungen, deren Dichten logarithmisch konkav sind, sind stets unimodal. Die Cauchy Verteilung ist aber ein Beispiel für eine unimodale Verteilung, deren Dichte nicht logarithmisch konkav ist. Diskrete Verteilungen erzeugen Treppenfunktionen als Verteilungsfunktionen und sind aus diesem Grunde mit Ausnahme der Einpunktverteilung alle multimodal. Für eine ausführliche Darstellung unimodaler Verteilungen verweisen wir auf das Buch von Dharmadhikari/Joag-dev (1988).

Beispiel: Die Abbildungen 3 und 4 zeigen die Verteilungsfunktionen bzw. Dichtefunktionen einer im Intervall $(0,1)$ gleichverteilten, einer standard-normalverteilten und einer U-förmig verteilten Zufallsvariablen.

Abbildung 3: Drei Verteilungsfunktionen (kumulativ normal, rechteckig, U-förmig)

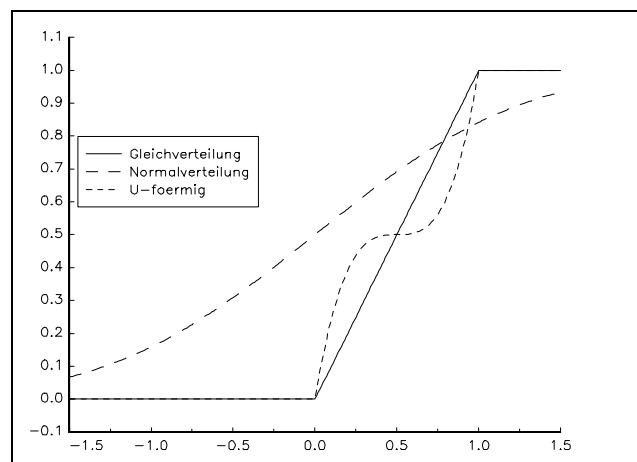
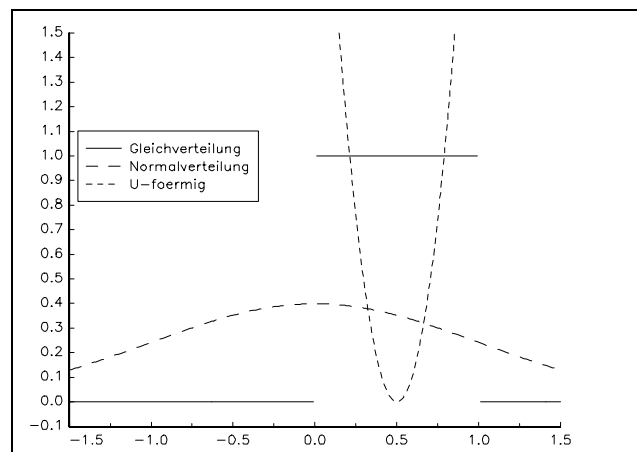


Abbildung 4: Dichtefunktionen für die Verteilungsfunktionen in Abbildung 3



Offensichtlich ist die Standard-Normalverteilung unimodal mit eindeutigem Modus 0, die Gleichverteilung unimodal, wobei für den Modus jeder Wert zwischen Null und Eins genommen werden kann, während die U-förmige Verteilung nicht unimodal, sondern multimodal ist. Beta- und Gamma-Verteilungen sind ebenfalls unimodal, wenn die Parameter größer gleich eins sind. Mischverteilungen sind weitere Beispiele, wie Multimodalität in der Praxis entstehen kann. Hat eine unimodale Verteilungsfunktion die Eigenschaft, daß die Faltung mit irgendeiner anderen unimodalen Verteilungsfunktion wieder zu einer unimodalen Verteilungsfunktion führt, so nennt man sie streng unimodal. Es läßt sich zeigen, daß streng unimodale nichtdegenerierte Verteilungsfunktionen genau solche sind, die stetig sind mit logarithmisch konkaven Dichten.

5. Der DIP-Test als Unimodalitätstest

Um den DIP-Test beschreiben zu können, brauchen wir ein Maß für den Abstand zweier Funktionen. Für zwei beschränkte Funktionen F und G sei ihr Abstand als $\rho(F,G)=\sup_x |F(x)-G(x)|$ definiert. Für eine Klasse \mathcal{A} von beschränkten Funktionen sei $\rho(F,\mathcal{A})=\inf_{G \in \mathcal{A}} \rho(F,G)$. Weiter sei \mathcal{U} die Klasse der unimodalen Verteilungsfunktionen. Der DIP einer Verteilungsfunktion F ist definiert durch $D(F)=\rho(F,\mathcal{U})$. Offensichtlich gilt $D(F)=0$ für $F \in \mathcal{U}$, $D(F)>0$ für $F \notin \mathcal{U}$ und aus der Dreiecksungleichung folgt $D(F_1) \leq D(F_2) + \rho(F_1,F_2)$. Daher mißt der DIP die Abweichung von der Unimodalität.

In der Praxis ist die theoretische Verteilung F in der Regel unbekannt. Was wir beobachten, ist eine Stichprobe X_1, X_2, \dots, X_n aus der Verteilung F . Bezeichnet F_n die empirische Verteilungsfunktion definiert durch $F_n(x) = (1/n) \sum \{X_i \leq x\}$, so folgt aus dem Glivenko-Cantelli'schen Theorem $\rho(F_n,F) \rightarrow 0$ fast überall und daher $D(F_n) \rightarrow D(F)$ fast überall. Ein Test, der auf dem DIP basiert, wird daher asymptotisch irgendeine unimodale Verteilungsfunktion von einer multimodalen unterscheiden.

- Zu testen ist: " H_0 : Verteilung ist unimodal" versus " H_1 : Verteilung ist multimodal".

Da die Nullhypothese zusammengesetzt ist, ist es bei der Entwicklung eines Tests notwendig, eine unimodale Verteilungsfunktion als Nullverteilung zu wählen. Hartigan/Hartigan (1985a) wählen die Gleichverteilung, weil sie vermuten, daß asymptotisch die DIP-Statistik stochastisch am größten für die Gleichverteilung ist. Ein DIP-Wert, der signifikant für diese Nullverteilung ist, wäre auch signifikant bezüglich jeder anderen unimodalen Verteilung. In diesem Sinn wäre die Gleichverteilung als schlechteste unimodale Verteilungsfunktion zu betrachten.

Die Berechnung des Dip-Wertes d erfolgt über einen Algorithmus, der eine unimodale Verteilungsfunktion konstruiert, die von F_n den Abstand d hat. Eine FORTRAN Subroutine wurde von Hartigan (1985b) veröffentlicht, die einen Fehler enthielt und von Sommer/McNamara (1987) korrigiert wurde.

Führen wir einen DIP-Test in den Beispielen durch, ergeben sich die in Abbildung 5 und 6 gezeigten Grafiken. Die Abbildungen zeigen die empirischen Verteilungsfunktionen nach unten und oben durch den Dip-Wert d verschoben. Die gestrichelte Linie gibt die gefittete unimodale Verteilungsfunktion an, die von der empirischen Verteilungsfunktion F_n gerade den Dip-Wert als Abstand hat. Sie besteht im Intervall $[x_1, x_L]$ aus der größten konvexen Minorante von $F_n + d$, im Intervall $[x_U, x_N]$ aus der kleinsten konkaven Majorante von $F_n - d$ und einem Liniensegment im Intervall (x_L, x_U) . Dabei ist x_1 der kleinste und x_N der größte x -Wert im Datensatz. x_L und x_U ergeben sich aus dem Algorithmus von Hartigan (1985b). In unserem Beispiel gilt jeweils $x_L = x_U$, was aber nicht immer so sein muß. Dieser Fall tritt aber insbesondere dann auf, wenn einige Ausprägungen stark dominieren, was bei Variablen mit wenigen Ausprägungen praktisch immer der Fall ist.

Abbildung 5: Verteilung 'Bedenken', nach oben und unten um den Dip-Wert d verschoben

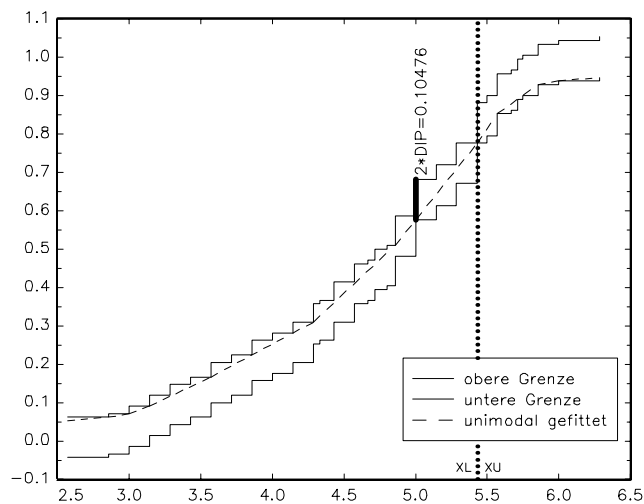
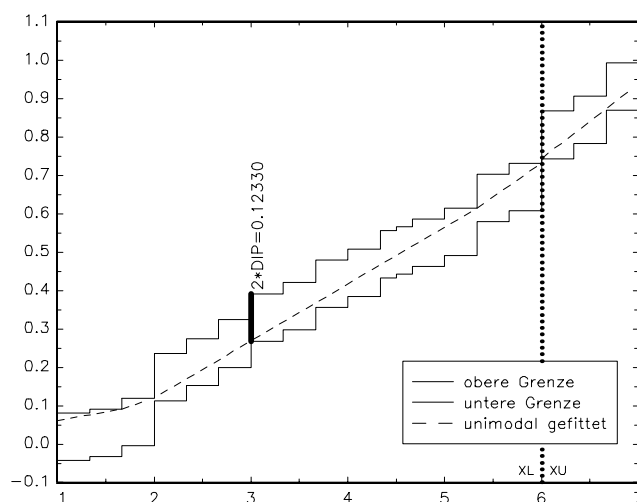


Abbildung 6: Verteilung 'Sorgen', nach oben und unten um den Dip-Wert d verschoben



Für die DIP-Statistik ergibt sich nach dem Algorithmus von Hartigan (1985b) für die Skalenwerte 'Bedenken' der Wert $d_B = 0.05238$, für 'Sorgen' der Wert $d_S = 0.06165$. Beide liegen nach der in Hartigan/Hartigan (1985a) und Hartigan (1985b) publizierte Tafel im Ablehnungsbereich eines 95% Konfidenzintervalls. Der Test besagt also, daß keiner der Datensätze aus eingipfligen Verteilungen stammt.

Bemerkenswert ist, daß im Falle 'Sorgen' $x_L = x_U$ ist und der berechnete Dip-Wert $d_S = 0.06165$ kleiner ist als $\delta = 7/103 = 0.067961$, wobei δ die relative Häufigkeit des Wertes x_L im Datensatz ist. Die gefittete unimodale Verteilungsfunktion ist daher an der Stelle x_L unstetig.

6. Diskussion

Zum DIP-Test ist zunächst anzumerken, daß dann, wenn die wahre unimodale Referenz-Verteilung stark von der Gleichverteilung abweicht, es andere Tests mit größerer Macht gibt (Hartigan/Hartigan1985a). Das sollte aber für den Anwender nicht von großer Bedeutung sein, weil diese wahre Verteilung im allgemeinen unbekannt sein dürfte. Man kann statt der Gleichverteilung auch die gefittete unimodale Verteilungsfunktion als

Nullverteilung verwenden. Simulationsstudien, die die Verteilung der DIP-Statistik bei Gültigkeit der neuen Nullverteilung sichtbar machen, ergaben aber in unseren Beispielen keine Änderung in der Entscheidung.

Lohnt sich der DIP-Test oder kommt man nicht in der Praxis mit dem in einigen Statistik-Paketen angebotenen Kolmogorov-Smirnov-Anpassungstest auf Normalität aus? Eine systematische Antwort läßt sich auf diese Frage zur Zeit nicht geben. Es ist aber bemerkenswert, wie das obige Anwendungsbeispiel zeigt, daß für die Variable 'Bedenken' der K-S Test die Normalverteilungshypothese beibehält, während der DIP-Test die schwächere Unimodalitätshypothese ablehnt. Die Entscheidungen fallen zwar knapp aus, aber dennoch zeigt sich der K-S Test zumindest in diesem Beispiel als der konservativere. Da beide Tests im Fall der Variablen 'Sorgen' jeweils zu einer Ablehnung der Nullhypothese führen, kann man sicher sagen, daß eine routinemäßige Anwendung des K-S Tests besser ist als die völlige Ignorierung der Unimodalitätsfrage. Dabei braucht man sich in den Programmpaketen nicht allein auf die Normalverteilung zu beschränken. Stand-alone Programme für den DIP-Test stehen aber ebenfalls zur Verfügung.

Wie häufig oder wie selten ist der Befund einer nicht-unimodalen Verteilung in der sozialwissenschaftlichen Praxis? Wir haben, um für diesen Artikel ein Anwendungsbeispiel zu finden, verschiedene Datensätze betrachtet. Dabei fanden wir z.B. für verschiedenste Meinungs- und Einstellungsitems des ALLBUS Unimodalität von Histogrammen. Eingipfeligkeit bei den Histogrammen für einzelne Items scheint also der dominierende Normalfall zu sein. Ob aber die den Stichproben zugrunde liegende Verteilung unimodal ist, läßt sich aus diesen Befunden nicht beantworten. Diese Frage wird vom DIP-Test - überraschend, wie es scheint - stets negativ beantwortet. Der Grund hierfür ist, daß im Falle der ALLBUS-Items stets nur eine Verteilung mit wenigen Skalenpunkten vorliegt. Das Kontinuum der X-Achse ist damit auf wenige Punkte "diskretisiert", d.h. die Werte klumpen, und der DIP-Test findet, ganz zu Recht, daß jeweils eine mehrgipfelige (in der Tat: eine viel-gipfelige) Verteilung angenommen werden muß. Im Fall der obigen Testvariablen 'Sorgen' und 'Bedenken' hatten wir dagegen einen anderen Fall, weil diese Variablen summative Werte aus mehreren Items darstellen, die viel feiner abgestuft sind. Die Frage, wie viele Abstufungen einer Variablen ausreichend sind, damit eine Stichprobenverteilung ausreichend viele Abstufungen enthält für eine sinnvolle Anwendung des DIP-Tests, ist zur Zeit noch offen. Eine denkbare praktische Lösung dieses Problems wäre vielleicht, die Verteilung zwischen den Skalenwerten einer grob abgestuften Skala zu interpolieren und dann hieraus eine Zufallsverteilung zu ziehen. In der Praxis sollte man aber, bevor solche Fragen mit einiger Sicherheit beantwortet werden können, bei

Variablen mit nur wenigen Abstufungen eher auf den χ^2 -Anpassungstest, den K-S oder den Lilliefors Test zurückgreifen.

Anmerkung

- 1) Eine Erweiterung auf mehr Dimensionen ist der SPAN-Test, beschrieben in Hartigan (1988).

Literatur

Borg, I., 1989: Zur Präsentation von Umfrageergebnissen. *Zeitschrift für Arbeits- und Organisationspsychologie* 33: 90-95.

Borg, I., 1992: Überlegungen und Untersuchungen zur Messung der subjektiven Unsicherheit der Arbeitsstelle. *Zeitschrift für Arbeits- und Organisationspsychologie* 36: 107-116.

Borg, I./Braun, M., 1992: Arbeitsethik und Arbeitsinvolvement als Moderatoren der psychologischen Auswirkungen von Arbeitsplatzunsicherheit. *Zeitschrift für Arbeits- und Organisationspsychologie* 36: 167-176.

Coombs, C. H., 1964: A theory of data. New York: Wiley.

Coombs, C. H./Avrunin, G., 1977: Single-peaked functions and the theory of preference. *Psychological Review* 84: 216-230.

Dharmadhikari, S./Joag-dev, K., 1988: Unimodality, Convexity, and Applications. Boston: Academic Press.

Hartigan, J.A./Hartigan, P.M., 1985a: The DIP-Test of unimodality. *The Annals of Statistics* 13: 70-84.

Hartigan, P.M., 1985b: Computation of the DIP-Statistic to test for unimodality. *Applied Statistics* 34: 320-325.

Hartigan, J.A., 1988: The SPAN test for unimodality. In H. H. Bock (Hrsg.) *Classification and Related Methods of Data Analysis*. Elsevier Science Publishers B.V. (North Holland).

Sommer, C.J./McNamara, J.N., 1987: Power considerations for the DIP-Test of unimodality using mixtures of normal and uniform distributions. *American Statistical Association; Proceedings of the Statistical Computing Section*: 186-191.

Thurstone, L.L., 1927: A law of comparative judgment. *Psychological Review* 34: 273-286.