

Sequentialverfahren

Prof. Dr. B. Schneider, Institut für Biometrie, Medizinische Hochschule Hannover

1 Grundlagen

Bei Sequentialverfahren ist kein fester Stichprobenumfang vorgegeben, bei dem eine statistische Analyse (Schätzung oder Test) ausgeführt werden soll; vielmehr kann bei der Datenerfassung (Beobachtung oder Experiment) laufend aufgrund der bisher erfaßten Daten entschieden werden, ob fortgesetzt oder abgeschlossen werden soll, wobei beim Abschluß die angestrebte Analyse mit allen erhobenen Daten ausgeführt wird. Im folgenden werden ausschließlich sequentielle Testverfahren behandelt. Für sequentielle Schätzverfahren sei auf das Buch von Bauer, Scheiber und Wohlzogen verwiesen [3].

Man kann den Verlauf eines Sequentialverfahrens grafisch durch einen Sequentialpfad in einem Koordinatensystem darstellen, bei dem auf der Abszisse der Stichprobenumfang (oder eine monotone Funktion davon) und auf der Ordinate die bei den jeweiligen Stichprobenumfängen aus den kumulierten Daten berechneten Werte einer Entscheidungsstatistik $t(x_1, \dots, x_m)$ aufgetragen sind. Zur Entscheidung über Fortsetzung oder Abschluß der Datenerhebung wird im Koordinatensystem ein **Fortsetzungsbereich** abgegrenzt. Solange sich der Pfad in diesem Bereich befindet, wird die Datenerhebung fortgesetzt, sobald der Pfad die Grenze des Bereichs erreicht, wird die entsprechende Analyse - bei statistischen Tests die entsprechende Testentscheidung - getroffen. Durch den Fortsetzungsbereich ist der **Sequentialplan** eindeutig festgelegt. In Abb. 1a und 1b sind solche Sequentialpläne (Sequentialpfade und Fortsetzungsbereiche) für eine einfache Testaufgabe dargestellt. Auf der Abszisse sind die Stichprobenumfänge $m=1, 2, \dots$, auf der Ordinate die Anzahlen r_m von Erfolgen aufgetragen, die bei den auf der Abszisse angegebenen Stichprobenumfängen m beobachtet wurden. Diese Ergebnisse können z. B. bei einer klinischen Phase II-Studie entstehen, bei der die therapeutische Wirksamkeit eines neuen Medikaments untersucht werden soll. Das Zielkriterium ist binär (0=kein Erfolg, 1=Erfolg). Das Medikament wird als aussichtsreich angesehen, wenn damit eine gute Erfolgswahrscheinlichkeit π (z. B. 60%) erreicht wird. Aufgrund der Studienergebnisse soll über die weitere Entwicklung entschieden werden. Wenn die Erfolgswahrscheinlichkeit π_0 (z. B. 40%) ist, soll eine Weiterentwicklung höchstens mit Wahrscheinlichkeit α (z. B. 5%) betrieben werden (Fehlerwahrscheinlichkeit 1. Art). Wenn die Erfolgswahrscheinlichkeit π_1 (z.B. 80%) ist, soll die Weiterentwicklung höchstens mit Wahrscheinlichkeit β (z. B. 20%) ausgeschlossen werden (Fehlerwahrscheinlichkeit 2. Art). Es ist also zwischen der Nullhypothese $H_0: \pi=\pi_0$ und der Alternative $H_1: \pi=\pi_1$ zu entscheiden. Bei Gültigkeit von H_0 soll die Nullhypothese höchstens mit der Irrtumswahrscheinlichkeit α abgelehnt, bei Gültigkeit von H_1 höchstens mit der Irrtumswahrscheinlichkeit β angenommen werden.

In Abb. 1 wird der Fortsetzungsbereich durch zwei parallele Geraden begrenzt. Es handelt sich um einen offenen Bereich. Man spricht daher auch von einem **offenen Sequentialplan**. In Abb. 2 wird der Fortsetzungsbereich durch einen geschlossenen Bereich (Dreieck) begrenzt. Dementsprechend nennt man diesen Sequentialplan **geschlossen**. Falls in beiden

Plänen der Sequentialpfad die obere Grenze des Fortsetzungsbereichs erreicht, wird die Nullhypothese abgelehnt und die Alternativhypothese angenommen, falls die untere Grenze erreicht wird, die Alternativhypothese abgelehnt und die Nullhypothese angenommen. Die obere Grenze des Fortsetzungsbereichs begrenzt somit auch den Annahmebereich für die Alternative, die untere den für die Nullhypothese.

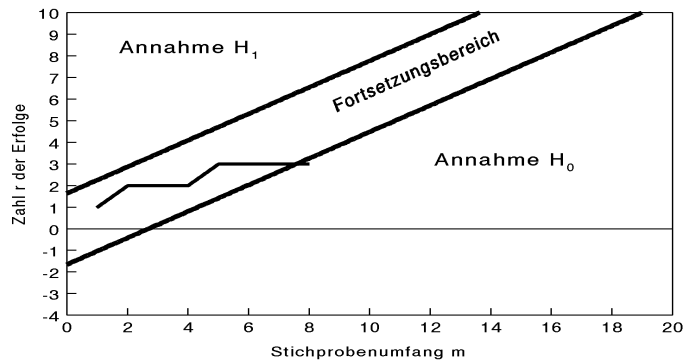


Abb.1 Offener Sequentialplan nach Wald zum Test von $\pi=\pi_0$ gegen $\pi=\pi_1$

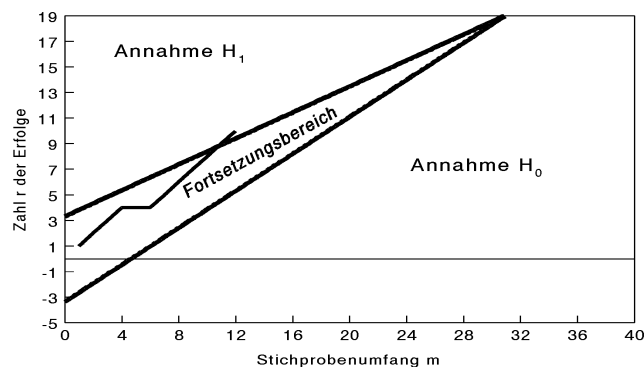


Abb. 2: Geschlossener Sequentialplan (Dreiecksplan) zum Test von $\pi=\pi_0$ gegen $\pi=\pi_1$

Die Wahrscheinlichkeit, daß der Pfad die obere bzw. untere Grenze des Fortsetzungsbereichs erreicht, hängt vom Wert des Parameters π ab. Die Wahrscheinlichkeit $P_0(\pi)$, die untere Grenze zu erreichen und so die Nullhypothese anzunehmen, wird bei Wald [6] auch die

‘Operating Characteristic Function’ (**OC-Funktion**) genannt. Die Wahrscheinlichkeit $P_1(\pi)$, die obere Grenze zu erreichen und damit die Alternative anzunehmen, ist die **Powerfunktion**. Bei sequentiellen Testplänen, bei denen die Pfade mit Wahrscheinlichkeit 1 eine der Grenzen erreichen (bei geschlossenen Plänen ist dies trivialerweise der Fall), gilt: $P_0(\pi)+P_1(\pi)=1$ für alle π . Es genügt daher zur Charakterisierung der Testgüte des Plans eine der beiden Funktionen. Wir werden im folgenden stets die Powerfunktion nehmen und sie mit $P(\pi)$ (bzw. allgemein bei einem Parameter θ mit $P(\theta)$) bezeichnen. Die OC-Funktion wird mit $OC(\pi)$ (allgemein $OC(\theta)$) bezeichnet.

Bei sequentiellen Testverfahren (Sequentialtests) sind die Grenzen des Fortsetzungsbereichs so zu wählen, daß $P(\pi_0)\leq\alpha$ und $P(\pi_1)\geq 1-\beta$ (bzw. $P(\theta_0)\leq\alpha$ und $P(\theta_1)\geq 1-\beta$, wenn allgemein θ_0 den durch die Nullhypothese und θ_1 den durch die Alternative festgelegten Parameterwert bezeichnen) gilt. Der Stichprobenumfang n , bei dem eine der beiden Entscheidungen getroffen wird, ist nun nicht mehr fest vorgegeben, sondern Realisation einer Zufallsgröße, deren Verteilung durch den Sequentialplan festgelegt ist und vom Parameterwert π (bzw. θ) abhängt. Zur Charakterisierung dieser Verteilung nimmt man üblicherweise den Erwartungswert $E(n|\pi)$ (bzw. $E(n|\theta)$). Dieser ist eine Funktion von π (bzw. θ) und wird als „Average Sample Number Function“ oder **ASN-Funktion** bezeichnet ($ASN(\theta)$). Allerdings kann die Verteilung von n oft sehr schief sein und somit der Erwartungswert die Verteilung nicht befriedigend charakterisieren. Besser wäre die Charakterisierung durch ausgezeichnete Quantilwerte (z. B. 25%-, 50%- und 75%-Quantile). Diese sind aber schwieriger zu berechnen.

Ein Sequentialplan, der die oben genannten Power-Forderungen zu gegebenem α und β einhält, ist um so besser je geringer für gegebene Parameterwerte seine ASN-Werte sind. Wir werden in den folgenden Abschnitten Sequentialpläne für wichtige Testprobleme angeben, die bei relevanten Parameterwerten minimale ASN-Werte haben. Diese können bis zur Hälfte der für dieselbe Testentscheidung bei festem Stichprobenumfang erforderlichen Anzahl betragen.

Die Festlegung der Annahmegrenzen (und damit des Sequentialplans) kann prinzipiell durch Simulation erfolgen: Man simuliert für die Parameterwerte der Nullhypothese und Alternative hinreichend viele Sequentialpfade und legt die Grenzen so fest, daß von den unter der Nullhypothese simulierten Pfaden höchstens der Anteil α die Grenze des Annahmebereichs für die Alternative erreicht, und von den unter der Alternative simulierten Pfaden höchstens der Anteil β die Grenze des Annahmebereichs für die Nullhypothese. Zu beachten ist, daß bereits das Erreichen einer Grenze durch den Sequentialpfad zum Abschluß des Verfahrens und Treffen der entsprechenden Entscheidung führt. Für die Entscheidung ist es irrelevant, ob der Pfad später wieder den Fortsetzungsbereich erreicht oder nicht. Die geometrische Form des Fortsetzungsbereichs kann im Prinzip beliebig sein. Man wird aber bestrebt sein, eine möglichst einfache Form zu erhalten. Die einfachste Form hat ein von Geraden begrenzter Bereich. Die Begrenzungsgeraden, bei denen die vorgegebenen Irrtumswahrscheinlichkeiten α und β eingehalten werden, lassen sich zumindest approximativ analytisch bestimmen. Wir werden im folgenden zwei Sequenzverfahren mit Fortsetzungsbereichen, die von Geraden begrenzt sind, besprechen: der Sequential Probability Ratio Test SPRT und Dreieckspläne

nach Whitehead [7, 8]. Die Bereiche dieser Pläne sind in Abb. 1 und 2 gezeigt.

Ein Sequentialpfad kann (bei unabhängigen Stichprobenwerten) als Realisation eines (diskreten) stochastischen Prozesses mit unabhängigen Zuwächsen (Irrfahrt-Prozeß, random walk process) aufgefaßt werden. Der Prozeß endet, wenn eine der Annahmegrenzen erreicht wird. Dies soll bei den durch die Hypothesen festgelegten Parameterwerten höchstens mit den vorgegebenen Irrtumswahrscheinlichkeiten erfolgen. Die Annahmegrenzen sind absorbierende Grenzen des Irrfahrt-Prozesses.

Es soll hier noch darauf hingewiesen werden, daß bei sequentiellen Verfahren die Entscheidung über Fortsetzung oder Abschluß nicht notwendig bei jeder neuen Beobachtung getroffen werden muß. Man kann die Beobachtungen in Gruppen zusammenfassen und erst nach den Ergebnissen einer Beobachtungsgruppe für Fortsetzung entscheiden, wenn der Pfad noch im Fortsetzungsbereich liegt, und für Annahme einer Hypothese, wenn der Pfad die entsprechende Annahmegrenze erreicht oder überschritten hat. Es ist keine bestimmte Gruppengröße vorzugeben; sie kann von Fall zu Fall variieren. Die Irrtumswahrscheinlichkeiten α und β werden bei diesem Vorgehen auf jedem Fall eingehalten. Es kann höchstens passieren, daß mehr Beobachtungen bis zur Entscheidung erhoben werden, als bei Einzelentscheidungen nötig wären, und damit die tatsächlichen Irrtumswahrscheinlichkeiten deutlich unter den vorgegebenen Werten liegen. Dieser als 'overshooting' bekannte Effekt ist allerdings bei nicht zu großer Gruppengröße gering. Er kann durch eine Korrektur der Annahmegrenzen für gruppensequentielle Entscheidungen noch weiter vermindert werden. Es bestehen also keine prinzipiellen Unterschiede zwischen „reinen“ Sequentialverfahren und „gruppensequentiellen“ Verfahren.

2 Der sequentielle Likelihood-Quotienten-Test (sequential probability ratio test SPRT)

Es wird folgendes Testproblem betrachtet: Die Beobachtungswerte x_1, x_2, \dots seien Realisationen identisch verteilter Zufallsgrößen mit der Verteilung $f(x;\theta)$, die den unbekanntem Parameter θ enthält. Dabei soll $f(x;\theta)$ für stetiges x die Verteilungsdichte und für diskretes (kategoriales) x die Wahrscheinlichkeit für den Wert x symbolisieren. Es soll die (einfache) Nullhypothese: $\theta=\theta_0$ gegen die Alternative: $\theta=\theta_1$ getestet werden, wobei die Nullhypothese höchstens mit der Irrtumswahrscheinlichkeit α und die Alternative höchstens mit der Irrtumswahrscheinlichkeit β abgelehnt werden soll.

A. Wald [6] hat hierfür das folgende sequentielle Testverfahren vorgeschlagen:

Als Entscheidungsstatistik wird die Folge der Likelihood-Quotienten:

$$LR_m = \frac{f(x_1, x_2, \dots, x_m; \theta_1)}{f(x_1, x_2, \dots, x_m; \theta_0)}$$

für $m=1, 2, \dots$ genommen und folgendermaßen entschieden:

- ◆ Fortsetzung der Beobachtungen, wenn: $B < LR_m < A$
- ◆ Entscheidung für Annahme der Alternative $\theta=\theta_1$, wenn: $LR_m \geq A$

◆ Entscheidung für Annahme der Nullhypothese $\theta = \theta_0$, wenn: $LR_m \leq B$

Die Entscheidungsgrenzen A und B sind so zu wählen, daß die vorgegebenen Irrtumswahrscheinlichkeiten α und β eingehalten werden. Dies ist der Fall, wenn folgende Grenzen genommen werden:

$$A = \frac{1-\beta}{\alpha} ; B = \frac{\beta}{1-\alpha}$$

Es ist hier anzumerken, daß diese Vorgaben von A und B konservativ sind. Streng genommen ist der Wert $(1-\beta)/\alpha$ eine obere Grenze für A und $\beta/(1-\alpha)$ eine untere Grenze für B. Dies kann durch folgende Überlegung eingesehen werden: Es wird für die Annahme der Alternative H_1 entschieden, sobald für ein m der Likelihoodquotient $LR_m \geq A$ ist. Alle Pfade, für die $f(x_1, \dots, x_m; \theta_1) \geq A \cdot f(x_1, \dots, x_m; \theta_0)$ für irgendein m gilt, führen zur Annahme von H_1 . Die Wahrscheinlichkeit, H_1 anzunehmen, ist somit bei Gültigkeit von H_1 mindestens A-mal so groß wie bei Gültigkeit von H_0 . Diese Wahrscheinlichkeit soll aber im ersten Fall mindestens $1-\beta$ und im zweiten Fall höchstens α sein. Daraus ergibt sich die Ungleichung: $1-\beta \geq A \cdot \alpha$ oder $A \leq (1-\beta)/\alpha$. Entsprechend führen alle Pfade, für die für irgendein m $f(x_1, \dots, x_m; \theta_1) \leq B \cdot f(x_1, \dots, x_m; \theta_0)$ gilt, zur Annahme von H_0 . Die Wahrscheinlichkeit, H_0 anzunehmen, ist somit bei Gültigkeit von H_0 mindestens B-mal so groß wie bei Gültigkeit von H_1 . Daraus folgt: $\beta \leq B \cdot (1-\alpha)$ oder $B \geq \beta/(1-\alpha)$. Wie Wald gezeigt hat, wären bei den exakten Grenzen (die genau α und β einhalten) zwischen 1% und 5% weniger Beobachtungen bis zu einer Entscheidung erforderlich.

Für die Powerfunktion $P(\theta)$ hat Wald folgende approximative Formel hergeleitet:

$$P(\theta) \approx \frac{1 - \left(\frac{\beta}{1-\alpha}\right)^h}{\left(\frac{1-\beta}{\alpha}\right)^h - \left(\frac{\beta}{1-\alpha}\right)^h} \quad \text{für } h \neq 0$$

wobei $h \neq 0$ eine Funktion von θ ist, die durch die Bedingung:

$$\int \left(\frac{f(x; \theta_1)}{f(x; \theta_0)} \right)^{h(\theta)} f(x; \theta) dx = 1$$

eindeutig als Funktion von θ bestimmt ist. Eine interessante Herleitung dieser Formel wurde von Wald in [6] gegeben, wobei vorausgesetzt wird, daß die Funktion $h(\theta)$ existiert und eindeutig ist (was im Anhang von [6] bewiesen wird). Es ist dann:

$$f^*(x; \theta) = \left(\frac{f(x; \theta_1)}{f(x; \theta_0)} \right)^{h(\theta)} f(x; \theta)$$

für $h \neq 0$ eine Verteilungsdichte. Zunächst sei $h(\theta) > 0$ angenommen und folgendes Testproblem betrachtet: Es soll die Hypothese $H^*_0: f(x; \theta)$ (d.h. $f(x; \theta)$ ist die richtige Verteilungsdichte) gegen die Alternative $H^*_1: f^*(x; \theta)$ (d.h. $f^*(x; \theta)$ ist die richtige Verteilungsdichte) geprüft

werden. Für diesen Test wird ein SPRT S^* mit folgendem Fortsetzungsbereich durchgeführt:

$$B^{h(\theta)} < \frac{f^*(x_1, \dots, x_m; \theta)}{f(x_1, \dots, x_m; \theta)} < A^{h(\theta)}$$

Da $f^*(x_1, \dots, x_m; \theta)/f(x_1, \dots, x_m; \theta) = (f(x_1, \dots, x_m; \theta_1)/f(x_1, \dots, x_m; \theta_0))^{h(\theta)}$ ist, entspricht dieser Test S^* genau dem SPRT S für $H_0: \theta = \theta_0$ gegen $H_1: \theta = \theta_1$ mit den Grenzen A und B . Es wird also genau dann H_0 bzw. H_1 mit S angenommen, wenn H^*_0 bzw. H^*_1 mit S^* angenommen wird. Bezeichnet $\alpha^*(\theta)$ die Wahrscheinlichkeit, mit S^* die Alternative H^*_1 anzunehmen, wenn H^*_0 gilt (d. h. $f(x; \theta)$ die Verteilungsdichte der Beobachtungen x_i ist), und $\beta^*(\theta)$ die Wahrscheinlichkeit, mit S^* die Nullhypothese H^*_0 anzunehmen, wenn H^*_1 gilt, dann ist (da S^* ein SPRT mit den Grenzen $A^{h(\theta)}$ und $B^{h(\theta)}$ ist):

$$A^{h(\theta)} \approx \frac{1 - \beta^*(\theta)}{\alpha^*(\theta)} \quad \text{und} \quad B^{h(\theta)} \approx \frac{\beta^*(\theta)}{1 - \alpha^*(\theta)}$$

Aus diesen beiden (approximativen) Gleichungen folgt:

$$\alpha^*(\theta) \approx \frac{1 - B^{h(\theta)}}{A^{h(\theta)} - B^{h(\theta)}} = \frac{1 - \left(\frac{\beta}{1 - \alpha}\right)^{h(\theta)}}{\left(\frac{1 - \beta}{\alpha}\right)^{h(\theta)} - \left(\frac{\beta}{1 - \alpha}\right)^{h(\theta)}}$$

Wird H^*_1 mit S^* angenommen, wenn H^*_0 gilt, dann wird H_1 mit S angenommen, wenn $f(x; \theta)$ die Verteilungsdichte der Beobachtungen x_i ist. Die Wahrscheinlichkeit dafür ist aber genau die Power $P(\theta)$. Es gilt also: $P(\theta) \approx \alpha^*(\theta)$.

Für $h < 0$ erhält man durch analoge Überlegungen dasselbe Ergebnis. Für $h \rightarrow 0$ ist: $P(\theta_{h=0}) = |\ln B| / (|\ln A| + |\ln B|)$.

Für die ASN-Funktion gilt folgende Formel:

$$ASN(\theta) \approx \frac{P(\theta) \ln A + (1 - P(\theta)) \ln B}{E_\theta(z)} \quad \text{für } E_\theta(z) \neq 0$$

wobei im Nenner der Erwartungswert (bei Parameterwert θ) von $Z = \ln(f(X; \theta_1)/f(X; \theta_0))$ steht und X die Zufallsgröße der Realisationen x_i symbolisiert. Die Gültigkeit dieser Formel kann folgendermaßen eingesehen werden:

Die Größen $z_i = \ln(f(x_i; \theta_1)/f(x_i; \theta_0))$ sind Realisationen von unabhängigen und identisch verteilten Zufallsgrößen Z_i mit dem Erwartungswert $E_\theta(z)$. Es sei n der Stichprobenumfang, bei dem das Verfahren zu einer Entscheidung kommt. n ist Realisation einer Zufallsgröße mit Erwartungswert $E_\theta(n) = ASN(\theta)$. Es sei N eine natürliche Zahl, für die $\Pr(n \geq N)$ vernachlässigt

werden kann. Aus $\sum_{i=1}^N Z_i = \sum_{i=1}^n Z_i + \sum_{j=n+1}^N Z_j$ folgt:

$$E_0\left(\sum_{i=1}^N Z_i\right) = NE_0(z) = E_0\left(\sum_{i=1}^n Z_i\right) + E_0\left(\sum_{j=n+1}^N Z_j\right).$$

Man beachte, daß die beiden letzten Erwartungswerte auch über n zu bilden sind. Für $j > n$ sind die Zufallsgrößen Z_j unabhängig von den Zufallsgrößen Z_i mit $i \leq n$. Es gilt daher:

$$E_0\left(\sum_{j=n+1}^N Z_j\right) = E_0(N - n)E_0(z) = NE_0(z) - E_0(n)E_0(z)$$

Eingesetzt in die vorhergehende Gleichung ergibt dies:

$$NE_0(z) = E_0\left(\sum_{i=1}^n Z_i\right) + NE_0(z) - E_0(n)E_0(z)$$

oder:

$$ASN(\theta) = E_0(n) = \frac{E_0\left(\sum_{i=1}^n Z_i\right)}{E_0(z)}$$

Nun ist (wenn 'overshooting' vernachlässigt wird) $\sum_{i=1}^n Z_i$ entweder $\ln A$ (mit Wahrscheinlichkeit $P(\theta)$) oder $\ln B$ (mit Wahrscheinlichkeit $1 - P(\theta)$), da bei n eine Entscheidung getroffen wird. Daraus folgt die oben angegebene Formel für $ASN(\theta)$.

Für θ^* mit $E_{\theta^*}(z) = 0$ gilt (wie Wald für stetiges $h(\theta)$ in [6] gezeigt hat) $h(\theta^*) = 0$ und

$$ASN(\theta^*) \approx \frac{|\ln A| \cdot |\ln B|}{E_{\theta^*}(z^2)} \text{ für } h(\theta^*) = 0 \text{ bzw. } E_{\theta^*}(z) = 0$$

Z.B. gilt für Bernoulli-verteilte Zufallsgrößen X mit $f(x=1; \theta) = \theta$ und $f(x=0; \theta) = 1 - \theta$:

$$\theta = \frac{1 - \left(\frac{1 - \theta_1}{1 - \theta_0}\right)^h}{\left(\frac{\theta_1}{\theta_0}\right)^h - \left(\frac{1 - \theta_1}{1 - \theta_0}\right)^h} \text{ für } h(\theta) \neq 0 \text{ und } \theta^* = \frac{\ln\left(\frac{1 - \theta_0}{1 - \theta_1}\right)}{\ln\left(\frac{\theta_1}{\theta_0}\right) - \ln\left(\frac{1 - \theta_1}{1 - \theta_0}\right)} \text{ für } h(\theta^*) = 0$$

$$E_0(z) = \theta \ln\left(\frac{\theta_1}{\theta_0}\right) + (1 - \theta) \ln\left(\frac{1 - \theta_1}{1 - \theta_0}\right); \quad E_{\theta^*}(z^2) = \ln\left(\frac{\theta_1}{\theta_0}\right) \cdot \ln\left(\frac{1 - \theta_0}{1 - \theta_1}\right)$$

Für normal verteilte Zufallsgrößen X mit Erwartungswert θ und Varianz σ^2 ist:

$$z = \frac{1}{2\sigma^2} (2(\theta_1 - \theta_0)x + \theta_0^2 - \theta_1^2)$$

$$h(\theta) = \frac{\theta_1 + \theta_0 - 2\theta}{\theta_1 - \theta_0}; \quad \theta^* = \frac{\theta_1 + \theta_0}{2};$$

$$E_0(z) = \frac{1}{2\sigma^2} (2(\theta_1 - \theta_0)\theta + \theta_0^2 - \theta_1^2); \quad E_{\theta^*}(z^2) = \frac{(\theta_1 - \theta_0)^2}{\sigma^2}$$

Wie Wald [6] gezeigt hat, besitzt der SPRT bei unabhängigen Stichprobenwerten x_i (für die $f(x_1, x_2, \dots, x_m; \theta) = f(x_1; \theta) f(x_2; \theta) \dots f(x_m; \theta)$ gilt) folgende Eigenschaften:

- ◆ Das Verfahren führt mit Wahrscheinlichkeit 1 nach endlich vielen Schritten zu einer Entscheidung (es ist entscheidungsdefinit).
- ◆ Für $\theta = \theta_0$ und $\theta = \theta_1$ besitzt das Verfahren unter allen vergleichbaren Sequenzverfahren minimale ASN.

3 SPRT für den 1-Stichprobenvergleich bei binären Daten x_i

Die Beobachtungen x_i sollen unabhängig sein und nur die Werte 1 (mit Wahrscheinlichkeit π) und 0 (mit Wahrscheinlichkeit $1-\pi$) annehmen. Zu testen ist die Hypothese: $\pi = \pi_0$ gegen die Alternative: $\pi = \pi_1$ mit vorgegebenen Irrtumswahrscheinlichkeiten α und β .

Nach m Beobachtungen soll r mal eine 1 und $m-r$ mal eine 0 festgestellt worden sein. Für den Likelihoodquotienten LR_m gilt: $LR_m = (\pi_1^r (1-\pi_1)^{m-r}) / (\pi_0^r (1-\pi_0)^{m-r})$. Durch Übergang zu Logarithmen erhält man folgenden Fortsetzungsbereich des SPRT:

$$\ln(B) < r \cdot \ln\left(\frac{\pi_1(1-\pi_0)}{\pi_0(1-\pi_1)}\right) + m \cdot \ln\left(\frac{1-\pi_1}{1-\pi_0}\right) < \ln(A)$$

Dies läßt sich umformen in:

$$a_0 + bm < r < a_1 + bm$$

Im $(r-m)$ -Koordinatensystem wird durch den linken bzw. rechten Ausdruck jeweils eine Gerade dargestellt, mit den Achsenabschnitten:

$$a_0 = \frac{\ln\left(\frac{\beta}{1-\alpha}\right)}{\ln\left(\frac{\pi_1(1-\pi_0)}{\pi_0(1-\pi_1)}\right)} \quad \text{bzw.} \quad a_1 = \frac{\ln\left(\frac{1-\beta}{\alpha}\right)}{\ln\left(\frac{\pi_1(1-\pi_0)}{\pi_0(1-\pi_1)}\right)}$$

und der Neigung:

$$b = \frac{\ln\left(\frac{1-\pi_0}{1-\pi_1}\right)}{\ln\left(\frac{\pi_1(1-\pi_0)}{\pi_0(1-\pi_1)}\right)}$$

In Abb. 1 sind diese parallelen Begrenzungsgeraden für $\pi_0=0.4$, $\pi_1=0.8$, $\alpha=0.05$ und $\beta=0.05$ gezeigt. Die Achsenabschnitte der beiden Geraden sind: $a_0=-1.643$ und $a_1=+1.643$; die (gemeinsame) Neigung ist: $b=0.613$. Die Entscheidungsregel lautet:

- ◆ Setze die Beobachtungen fort, solange $a_0 + bm < r < a_1 + bm$.

- ◆ Entscheide für $\pi=\pi_0$, wenn $r \leq a_0 + bm$.
- ◆ Entscheide für $\pi=\pi_1$, wenn $r \geq a_1 + bm$.

Die Powerfunktion $P(\pi)$ dieses Verfahrens ist in Abb. 3 gezeigt. Die Formel lautet:

$$P(\pi) = \frac{1 - \left(\frac{\beta}{1-\alpha}\right)^h}{\left(\frac{1-\beta}{\alpha}\right)^h - \left(\frac{\beta}{1-\alpha}\right)^h}$$

wobei h durch:

$$h = \frac{1 - \left(\frac{1-\pi_1}{1-\pi_0}\right)^h}{\left(\frac{\pi_1}{\pi_0}\right)^h - \left(\frac{1-\pi_1}{1-\pi_0}\right)^h}$$

eindeutig (für $h \neq 0$) von π abhängt. Dem Wert $h=0$ entspricht $\pi=b$ und $P(b)=|a_0|/(a_1+|a_0|)$.

Für die ASN-Funktion gilt die Formel (für $\pi \neq b$ entsprechend $h(\pi) \neq 0$):

$$ASN(\pi) = \frac{P(\pi) \ln A + (1 - P(\pi)) \ln B}{\pi \ln \frac{\pi_1}{\pi_0} + (1 - \pi) \ln \frac{1 - \pi_1}{1 - \pi_0}}$$

Für $\pi=b$ ($h=0$) ist ASN:

$$ASN(b) = \frac{|\ln A| \cdot |\ln B|}{\ln(\pi_1 / \pi_0) \cdot \ln((1 - \pi_0)/(1 - \pi_1))}.$$

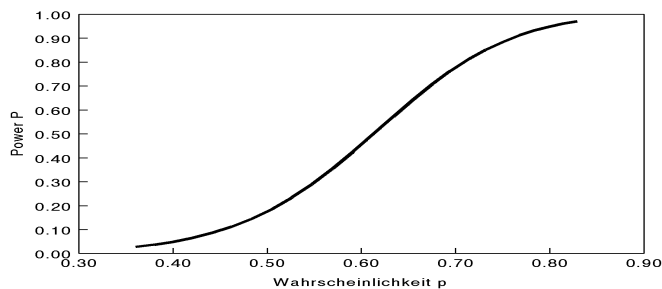


Abb. 3 Powerfunktion des offenen Sequentialplans (von Abb.1)

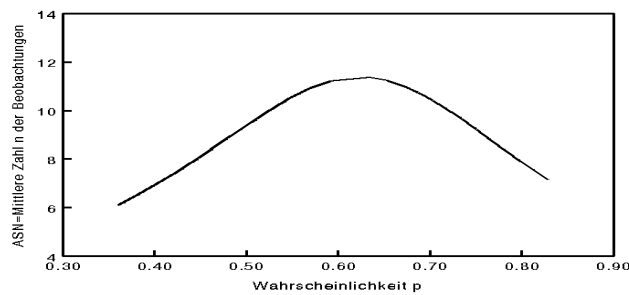


Abb. 4 ASN-Funktion

Für die Beispieldaten ist diese Funktion in Abb. 4 gezeigt. Im Mittel wird bei $\pi=\pi_0$ das Sequenzverfahren nach $n=7$ und bei $\pi=\pi_1$ nach $n=8$ Beobachtungen beendet. Bei fest vorgegebenem Stichprobenumfang sind approximativ:

$$n_{\text{fix}} = \frac{(z_{1-\alpha} \sqrt{\pi_0(1-\pi_0)} + z_{1-\beta} \sqrt{\pi_1(1-\pi_1)})^2}{(\pi_1 - \pi_0)^2}$$

Beobachtungen erforderlich, um die Irrtumswahrscheinlichkeiten α und β einzuhalten. Dies ergibt für das Beispiel: $n_{\text{fix}}=14$ Beobachtungen. Die mittlere Anzahl der Beobachtungen ist also beim SPRT etwa halb so groß wie beim festen Stichprobenplan, wenn die Nullhypothese oder Alternative zutrifft.

4 Anwendung auf 2-Stichprobenvergleiche

Bei vielen praktischen Problemen soll entschieden werden, ob der Wert eines Zielparameters bei zwei Stichproben gleich ist, oder ob der Zielparameter in der einen Stichprobe einen größeren (oder kleineren) Wert hat als in der anderen. So kann bei einer klinischen Studie danach gefragt werden, ob die Erfolgswahrscheinlichkeit π_1 einer Prüfbehandlung gleich der Erfolgswahrscheinlichkeit π_2 einer Standardbehandlung ist, oder ob $\pi_1 > \pi_2$ gilt; oder es soll entschieden werden, ob die Erwartungswerte einer Zielgröße (z. B. Blutdrucksenkung) für die Prüf- und Standardbehandlung gleich sind, oder ob der Erwartungswert für die Prüfbehandlung größer als für die Standardbehandlung ist. Die sequentielle Testung dieser Hypothesen kann auf das im Abschnitt 3 geschilderte Verfahren zurückgeführt werden, wenn jeweils Paare von beiden Stichproben (z. B. ein mit dem Prüfmittel und ein mit dem Standardmittel behandelter Patient) sequentiell beobachtet werden.

4.1 Paarweiser Vergleich binärer Ereignisse

Es sollen die Wahrscheinlichkeiten π_1 und π_2 zweier unabhängiger, Bernoulli-verteilter Zufallsgrößen verglichen werden. Das Problem tritt z. B. bei einer komparativen klinischen Studie auf, in der die Erfolgswahrscheinlichkeiten zweier Behandlungen A und B verglichen werden sollen. Dabei wird jeweils von einem Patientenpaar der eine mit A und der andere mit B behandelt (bei randomisierter Zuteilung) und es werden die Erfolge und Mißerfolge der beiden Behandlungen beobachtet. Zu entscheiden ist zwischen der Nullhypothese $H_0: \pi_1 = \pi_2$ und der Alternative $H_1: \pi_1 > \pi_2$.

Wird ein Erfolg mit 1 und ein Mißerfolg mit 0 kodiert, dann sind bei einem Paar die 4 Kombinationen: (1,1) (1,0) (0,1) (0,0) möglich, wobei die erste Zahl des Ergebnis unter A und die zweite das unter B angibt. Die beiden konkordanten Kombinationen (1,1) und (0,0) tragen nicht zur Unterscheidung der beiden Hypothesen bei und werden daher nicht berücksichtigt. Es wird somit im Sequenzverfahren nur die Folge der diskordanten Paare (1,0) und (0,1) berücksichtigt. In dieser Folge ist die Wahrscheinlichkeit für ein Paar (1,0)

$$\Pi = \frac{\pi_1(1 - \pi_2)}{\pi_1(1 - \pi_2) + \pi_2(1 - \pi_1)}$$

und die Wahrscheinlichkeit für ein Paar (0,1) $1 - \Pi$. Die Nullhypothese H_0 ist somit äquivalent der Hypothese $H'_0: \Pi = \Pi_0 = 1/2$. Als Alternative sind zwei Werte π_1 und π_2 ($\pi_1 > \pi_2$) vorzugeben, deren Unterschied als relevant für die Erfolgswahrscheinlichkeiten angesehen wird. Die Alternative H'_1 nimmt für Π den Wert Π_1 an, der sich daraus ergibt. Der 2-Stichprobenvergleich wird so auf einen 1-Stichprobenvergleich der beiden Hypothesen H'_0 und H'_1 zurückgeführt, der sequentiell mit dem SPRT des Abschnitts 3 durchgeführt werden kann. Die Zahl m von Beobachtungen ist dabei die Zahl der beobachteten diskordanten Paare, r die Zahl der dabei beobachteten Paare (1,0). Falls die Alternative $\pi_1 < \pi_2$ geprüft werden soll, ist die Zahl r der Paare (0,1) sequentiell über der Zahl m der diskordanten Paare zu nehmen und Π_1 entsprechend zu definieren.

Wie Wald [6] gezeigt hat, wird durch die Beschränkung auf die diskordanten Paare der ASN-Wert nicht wesentlich vergrößert. Diese Beschränkung erlaubt aber die Festlegung der modifizierten Nullhypothese auf den Wert $1/2$.

4.2 Paarweiser Vergleich von zwei quantitativen oder ordinalen Meßwerten

Auch der Vergleich von zwei quantitativen oder ordinalen Zufallsgrößen X_1 und X_2 kann durch Paarbildung auf den 1-Stichproben SPRT für binäre Variablen zurückgeführt werden. Bei den sequentiell beobachteten Meßwertpaaren (x_1, x_2) wird jeweils festgestellt, ob $x_1 > x_2$ (+Paar) oder $x_1 < x_2$ (-Paar) ist. Paare, bei denen dies nicht entschieden werden kann, werden nicht berücksichtigt. In der modifizierten Nullhypothese H_0 wird festgelegt, daß die Wahrscheinlichkeit für ein +Paar gleich der Wahrscheinlichkeit für ein -Paar ist; d. h. $\pi_0 = 1/2$, wobei $\pi = P(X_1 > X_2)$ die Wahrscheinlichkeit für ein +Paar sein soll. Für die Alternative ist ein geeigneter Wert π_1 festzulegen. Unter der Annahme, daß X_1 und X_2 zwei unabhängige,

normal verteilte Zufallsgrößen mit den Mittelwerten μ_1 und μ_2 und einer gemeinsamen Standardabweichung σ sind, gilt: $\pi = \Phi(\delta \cdot 2/\sigma)$, wobei $\delta = \mu_1 - \mu_2$ und Φ die Standard-Normalverteilungsfunktion bedeuten. Bei Vorgabe einer relevanten Differenz δ_1 kann der entsprechende Wert π_1 berechnet werden. Zu diesen beiden Wahrscheinlichkeiten $\pi_0 (=1/2)$ und π_1 und zu vorgegebenem α und β ist der SPRT zu konstruieren. Die Zahl r der +Paare ist über die Zahl m der unterschiedlichen Paare aufzutragen.

4.3 Der sequentielle t-Test

Beim Test der Unterschiede in den Mittelwerten μ_1 und μ_2 zweier unabhängiger und normal verteilter Zufallsgrößen X_1 und X_2 mit gleicher Standardabweichung σ hängt die Likelihood nicht nur von der zu testenden Differenz $\mu_1 - \mu_2$, sondern auch noch vom Störparameter σ ab. Die Likelihoodquotienten sind daher Funktionen des Störparameters σ . Bei dem oben angegebenen Verfahren (das auf Armitage [1] zurückgeht), wird diese Abhängigkeit durch die Transformation: $z_i = \text{sign}(x_{1i} - x_{2i})$ aufgehoben. Allgemein hat Wald [6] vorgeschlagen, diese Schwierigkeit durch die Einführung von Gewichtsfunktionen für die Störparameter zu beseitigen. Die Likelihoodquotienten des SPRT sind dann die Quotienten der gewichteten Likelihoodfunktionen unter Alternativ- und Nullhypothese. Eine andere Möglichkeit besteht darin, statt der Meßwertfolge eine Folge von daraus berechneten Statistiken zu nehmen, deren Verteilung von den Störparametern unabhängig ist. Beim Vergleich der Erwartungswerte zweier normal verteilter Zufallsgrößen mit gleicher Varianz erfüllt die t-Statistik diese Bedingung. Statt der Folge der ursprünglichen Meßwerte wird die Folge t_1, t_2, \dots der t-Statistiken betrachtet. Zu testen ist die Nullhypothese $H_0: \delta = (\mu_1 - \mu_2)/\sigma = 0$ gegen die Alternative $H_1: \delta = \delta_1 (\neq 0)$. Für den SPRT ist die zur Folge der t-Statistiken gehörende Folge der Likelihoodquotienten LR_m zu bilden und das Sequenzverfahren mit der Annahme der Nullhypothese zu beenden, wenn für ein m die Größe $LR_m \leq B$ ist, und mit der Annahme der Alternativhypothese, wenn $LR_m \geq A$ ist, wobei $B = \beta/(1-\alpha)$ und $A = (1-\beta)/\alpha$ sind.

Die t-Statistik hat unter H_0 eine zentrale und unter H_1 eine nichtzentrale t-Verteilung mit dem Nichtzentralitätsparameter $\delta_1(n_1 n_2 / (n_1 + n_2))^{1/2}$, wenn die t-Statistik aus n_1 Stichprobenwerten der ersten und n_2 der zweiten Stichprobe berechnet wird. Es müssen somit für $m=1, 2, \dots$ die Werte der Dichten $f(t_1, \dots, t_m; \delta)$ für $\delta = \delta_1$ und $\delta = 0$ berechnet und der Quotient gebildet werden. Dies kann durch die Faktorisierung: $f(t_1, \dots, t_m; \delta) = f(t_m; \delta) f(t_1, \dots, t_{m-1} | t_m)$ vereinfacht werden, da der zweite Faktor nicht mehr von δ abhängt und deshalb bei der Quotientenbildung entfällt. Die Folge der Likelihoodquotienten LR_m hat somit die Form: $f(t_m; \delta_1)/f(t_m; 0)$, wobei im Zähler die Dichte der nichtzentralen t-Verteilung mit dem von δ_1 und den Stichprobenumfängen abhängigen Nichtzentralitätsparameter und im Nenner die Dichte der zentralen t-Verteilung stehen. Formeln für diese Dichten sind bei Wald [6] und Wetherill [7] angegeben. Zur Berechnung der t-Statistiken sind jeweils fortlaufende Meßwerte aus beiden Stichproben zu nehmen. Man kann diese Stichproben in Gruppen vom Umfang $k (>1)$ einteilen. Zur Berechnung von t_1 werden die Werte x_{11}, \dots, x_{1k} der ersten Stichprobe und die Werte x_{21}, \dots, x_{2k} der zweiten Stichprobe verwendet, zur Berechnung von t_2 zusätzlich die nächsten k Werte beider Stichproben usw. Der Likelihoodquotient zur Statistik t_m ist dann der Quotient aus dem Dichtewert $f(t_m; \delta_1)$ der nichtzentralen t-Verteilung mit $k \cdot m - 2$ Freiheitsgraden und dem Nichtzentralitätsparameter $\delta_1(k \cdot m / 2)^{1/2}$ und dem Wert $f(t_m)$ der zentralen t-Verteilung mit $k \cdot m -$

2 Freiheitsgraden.

Nach diesem Prinzip lassen sich auch mit anderen Teststatistiken (Chi²-Test, F-Test) sequentielle Testverfahren zu den entsprechenden Hypothesen konstruieren, soweit sich die Verteilungsdichten dieser Statistiken faktorisieren lassen. Bedingungen für diese Faktorisierung sind in einer Arbeit von Cox [4] angegeben. Ein Problem dürfte allerdings die Verfügbarkeit geeigneter Algorithmen oder Programme zur Berechnung der Dichteverteilung sein. Deshalb wurden bisher diese sequentiellen Tests kaum angewandt. Im nächsten Abschnitt werden eine Approximation der Likelihoodquotienten und darauf beruhende Sequenzverfahren angegeben, mit denen für eine Vielzahl von Testproblemen sequentielle Verfahren leicht konstruiert werden können.

5 Approximation des Log-Likelihoodquotienten mit den Statistiken Z und V

Wir nehmen an, daß für einen Parameter θ der Verteilungsdichte $f(x;\theta)$ die Nullhypothese $\theta=0$ gegen eine Alternative $\theta=\theta_1$ mit den Stichprobenwerten x_1, \dots, x_m getestet werden soll. Die Stichprobenwerte sind Realisationen von unabhängigen und identisch verteilten Zufallsgrößen mit der Verteilungsdichte $f(x;\theta)$. Der Logarithmus der Likelihood: $l(\mathbf{x};\theta)=\ln(f(\mathbf{x};\theta))$ hängt für gegebene \mathbf{x} -Werte (die mit \mathbf{x} symbolisiert werden) vom Parameter θ (und u. U. noch von Störparametern ψ) ab. Ist nur der Parameter θ vorhanden, dann kann $l(\mathbf{x};\theta)$ für θ -Werte, die nicht allzuweit von 0 entfernt sind, in eine Taylorreihe um $\theta=0$ entwickelt werden:

$$l(\mathbf{x};\theta) = l(\mathbf{x};0) + \theta l_{\theta}(\mathbf{x};0) + \frac{1}{2}\theta^2 l_{\theta\theta}(\mathbf{x};0) + O(\theta^3)$$

Dabei bedeuten l_{θ} die erste Ableitung und $l_{\theta\theta}$ die zweite Ableitung nach θ .

Mit der Bezeichnung:

$$Z = l_{\theta}(\mathbf{x};0) \quad \text{und} \quad V = -l_{\theta\theta}(\mathbf{x};0)$$

und unter Vernachlässigung von Gliedern der Ordnung $O(\theta^3)$ lautet somit die Log-Likelihood:

$$l(\mathbf{x};\theta) = \text{const.} + \theta Z - \frac{1}{2}\theta^2 V$$

Hängt die Likelihood-Funktion noch von einem Vektor von Störparametern $\psi=(\psi_1, \psi_2, \dots, \psi_k)'$ ab, so sind diese geeignet zu ersetzen. Entsprechend dem Vorschlag in Whitehead [7, 8] werden in der Log-Likelihood-Funktion die Störparameter durch ihre Maximum-Likelihood-Schätzwerte für gegebenes θ ersetzt. Damit ist die Log-Likelihoodfunktion nur noch eine Funktion von θ , die um $\theta=0$ in eine Taylorreihe entwickelt wird. Diese Entwicklung wird nun hergeleitet. Zunächst werden einige Hilfsformeln gebracht:

Maximum-Likelihood-Schätzer (ML-Schätzer) für ψ bei gegebenem θ (die mit $\psi^{\wedge}(\theta)$ bezeichnet werden) sind Lösungen des nichtlinearen Gleichungssystems: $\mathbf{I}_{\psi}(\theta, \psi)=0$. Dabei

bedeutet \mathbf{l}_ψ den Vektor der Ableitungen von l nach ψ_i . Der ML-Schätzer bei $\theta=0$ wird mit ψ^* ($=\psi^\wedge(0)$) bezeichnet. Durch Taylorentwicklung von $\psi^\wedge(\theta)$ um $\theta=0$ erhält man:

$$\psi^\wedge(\theta) = \psi^* + \theta \psi^\wedge_\theta(0) + O(\theta^2)$$

wobei ψ^\wedge_θ der Vektor der Ableitungen von $\psi^\wedge(\theta)$ nach θ ist. Dieser Vektor kann als Lösung eines linearen Gleichungssystems mit der Matrix der zweiten Ableitungen von $l(\theta, \psi)$ nach ψ_i und ψ_j als Koeffizienten dargestellt werden, die als $\mathbf{l}_{\psi\psi}(\theta, \psi^\wedge(\theta))$ geschrieben wird. Aus der ML-Gleichung $\mathbf{l}_\psi(\theta, \psi^\wedge(\theta))=0$ folgt nämlich:

$$\frac{d}{dt} (\mathbf{l}_\psi(\theta, \psi^\wedge(\theta))) = \mathbf{l}_{\theta\psi}(\theta, \psi^\wedge(\theta)) + \mathbf{l}_{\psi\psi}(\theta, \psi^\wedge(\theta)) \psi^\wedge_\theta(\theta) = 0$$

und daraus:

$$\psi^\wedge_\theta(\theta) = -\mathbf{l}_{\psi\psi}^{-1}(\theta, \psi^\wedge(\theta)) \mathbf{l}_{\theta\psi}(\theta, \psi^\wedge(\theta))$$

Insbesondere gilt:

$$\psi^\wedge_\theta(0) = -\mathbf{l}_{\psi\psi}^{-1}(0, \psi^*) \mathbf{l}_{\theta\psi}(0, \psi^*)$$

und

$$\psi^\wedge(\theta) - \psi^* = -\theta \mathbf{l}_{\psi\psi}^{-1}(0, \psi^*) \mathbf{l}_{\theta\psi}(0, \psi^*) + O(\theta^2)$$

Es wird nun die Taylorentwicklung von $l(\theta, \psi^\wedge(\theta))$ um $(0, \psi^*)$ betrachtet:

$$l(\theta, \psi^\wedge(\theta)) = l(0, \psi^*) + \theta \mathbf{l}_{\theta}(0, \psi^*) + (\psi^\wedge(\theta) - \psi^*)' \mathbf{l}_\psi(0, \psi^*) + \frac{1}{2} \theta^2 \mathbf{l}_{\theta\theta}(0, \psi^*) + \theta \mathbf{l}_{\theta\psi}(0, \psi^*)' (\psi^\wedge(\theta) - \psi^*) + \frac{1}{2} (\psi^\wedge(\theta) - \psi^*)' \mathbf{l}_{\psi\psi}(0, \psi^*) (\psi^\wedge(\theta) - \psi^*) + O(\theta^3)$$

Das Glied $(\psi^\wedge(\theta) - \psi^*)' \mathbf{l}_\psi(0, \psi^*)$ ist 0, da ψ^* ML-Schätzer bei $\theta=0$ ist.

In den beiden letzten Gliedern wird $\psi^\wedge(\theta) - \psi^*$ durch $-\theta \mathbf{l}_{\psi\psi}^{-1}(0, \psi^*) \mathbf{l}_{\theta\psi}(0, \psi^*)$ ersetzt. Dies ergibt:

$$\theta \mathbf{l}_{\theta\psi}(0, \psi^*)' (\psi^\wedge(\theta) - \psi^*) = -\theta^2 \mathbf{l}_{\theta\psi}(0, \psi^*)' \mathbf{l}_{\psi\psi}^{-1}(0, \psi^*) \mathbf{l}_{\theta\psi}(0, \psi^*)$$

und

$$\frac{1}{2} (\psi^\wedge(\theta) - \psi^*)' \mathbf{l}_{\psi\psi}(0, \psi^*) (\psi^\wedge(\theta) - \psi^*) = \frac{1}{2} \theta^2 \mathbf{l}_{\theta\psi}(0, \psi^*)' \mathbf{l}_{\psi\psi}^{-1}(0, \psi^*) \mathbf{l}_{\theta\psi}(0, \psi^*)$$

Die Summe dieser beiden Glieder ist: $-\frac{1}{2} \theta^2 \mathbf{l}_{\theta\psi}(0, \psi^*)' \mathbf{l}_{\psi\psi}^{-1}(0, \psi^*) \mathbf{l}_{\theta\psi}(0, \psi^*)$.

Dies ergibt für die Taylorentwicklung:

$$l(\theta, \psi^\wedge(\theta)) = l(0, \psi^*) + \theta \mathbf{l}_{\theta}(0, \psi^*) + \frac{1}{2} \theta^2 (\mathbf{l}_{\theta\theta}(0, \psi^*) - \mathbf{l}_{\theta\psi}(0, \psi^*)' \mathbf{l}_{\psi\psi}^{-1}(0, \psi^*) \mathbf{l}_{\theta\psi}(0, \psi^*)) + O(\theta^3) \\ = \text{const.} + \theta Z - \frac{1}{2} \theta^2 V + O(\theta^3).$$

mit:

$$Z = \mathbf{l}_{\theta}(0, \psi^*)$$

$$V = -(\mathbf{l}_{\theta\theta}(0, \psi^*) - \mathbf{l}_{\theta\psi}(0, \psi^*)' \mathbf{l}_{\psi\psi}^{-1}(0, \psi^*) \mathbf{l}_{\theta\psi}(0, \psi^*))$$

Im folgenden wird angenommen, daß die Verteilungsdichte $f(x; \theta)$ nur vom Parameter θ abhängt. Falls Störparameter vorhanden sind, dann sollen diese durch ihre Maximum-

Likelihood-Schätzer bei gegebenem θ ersetzt sein.

Die Statistik $Z=l_{\theta}(\mathbf{x};0)=\sum_i \partial \ln(f(x_i;0)/\partial \theta)=\sum_i f_{\theta}(x_i;0)/f(x_i;0)$ wird 'Score-Statistik' (efficient score) genannt. Sie charakterisiert Abweichungen von der Nullhypothese. Die Statistik $V=-l_{\theta\theta}(\mathbf{x};0)$ hängt mit 'Fisher's Information' $I(\theta)$ zusammen. Diese ist definiert als:

$$I(\theta)=E((l_{\theta}(\mathbf{X};\theta))^2|\theta)=-E(l_{\theta\theta}(\mathbf{X};\theta)|\theta)=-\sum_i E(\partial^2 \ln(f(X_i;\theta))/\partial \theta^2|\theta)=m \cdot i(\theta)$$

wobei die Stichprobenwerte x_1, \dots, x_m durch die entsprechenden Zufallsgrößen X_i ersetzt sind und $E(\cdot|\theta)$ den Erwartungswert bezüglich der Verteilung $f(x;\theta)$ symbolisiert. Die Größe $i(\theta)=E((\partial \ln(f(X;\theta))/\partial \theta)^2|\theta)=-E(\partial^2 \ln(f(X;\theta))/\partial \theta^2|\theta)$ ist die Information einer Einzelbeobachtung bezüglich des Parameters θ .

Die Statistik V ist ein Schätzwert für $I(0)=m \cdot i(0)$, der für $\theta=0$ erwartungstreu ist; d. h. $E(V|0)=I(0)=m \cdot i(0)$. Sie kann als Kennwert des 'informativen' Stichprobenumfangs angesehen werden, der bezüglich θ in Z enthalten ist. Asymptotisch strebt V für $\theta=0$ gegen $I(0)$ und kann für nicht zu große Beträge $|\theta|$ als Approximation von $I(\theta)$ angesehen werden. Die Approximation gilt exakt, wenn $l_{\theta\theta}$ nicht von den Stichprobenwerten und θ abhängt. Dies ist z.B. bei m Stichprobenwerten der Fall, die unabhängige Realisationen einer normal verteilten Zufallsgröße mit Mittelwert θ und Varianz 1 sind. Es ist dann: $l_{\theta}=\sum x_i - m\theta$ und $l_{\theta\theta}=-m$; d.h. $Z=\sum x_i$ und $V=m=I(0)=I(\theta)$, da I nicht von θ abhängt.

Die Statistik $Z=l_{\theta}(0)=\sum_i \partial \ln(f(X_i;0))/\partial \theta=\sum_i f_{\theta}(X_i;0)/f(X_i;0)$ ist als Summe von unabhängig und identisch verteilten Zufallsgrößen nach dem Hauptgrenzwertsatz asymptotisch normalverteilt. Aus der Approximation: $l_{\theta}(\theta) \approx l_{\theta}(0) + \theta l_{\theta\theta}(0)$ folgt:

$$E(Z|\theta)=E(l_{\theta}(0)|\theta) \approx E(l_{\theta}(\theta)|\theta) - \theta E(l_{\theta\theta}(0)|\theta)$$

Bei regulären Likelihoodfunktionen (bei denen Integration und Differentiation vertauscht werden können) gilt:

$$E(l_{\theta}(\theta)|\theta)=\int (\partial \ln(f(x;\theta)/\partial \theta) f(x;\theta) dx = \int (f_{\theta}(x;\theta)/f(x;\theta)) f(x;\theta) dx = \partial / \partial \theta (\int f(x;\theta) dx) = 0$$

da das letzte Integral unabhängig von θ gleich 1 ist. Es gilt somit approximativ:

$$E(Z|\theta) \approx -\theta E(l_{\theta\theta}(0)|\theta) \approx \theta V$$

wobei im letzten Ausdruck der Erwartungswert $-E(l_{\theta\theta}(0)|\theta)$ durch den Schätzwert V ersetzt wurde.

Für die Varianz von $Z=l_{\theta}(0)$ gilt:

$$\text{var}(Z) = E(l_{\theta}(0)^2|\theta) - \theta^2 E(l_{\theta\theta}(0)|\theta)^2 = E(l_{\theta\theta}(0)|\theta)(1 - \theta^2 E(l_{\theta\theta}(0)|\theta)) \approx V(1 - \theta^2 V) \approx V$$

wenn $|\theta|$ hinreichend klein ist, so daß $\theta^2 V$ vernachlässigt werden kann.

Approximativ gilt somit für hinreichend große Stichprobenumfänge m und kleine $|\theta|$ (d. h. bis auf Ausdrücke der Ordnung $O(\theta^2)$):

$$Z \sim N(\theta V, \sqrt{V}) ;$$

d.h. Z ist asymptotisch normal verteilt mit Mittelwert $-\theta E(l_{\theta\theta}(0)|\theta) \approx \theta V$ und Varianz $E(l_{\theta\theta}(0)|\theta)(1-\theta^2 E(l_{\theta\theta}(0)|\theta)) \approx V$.

Um mit Z die Hypothese $\theta=0$ gegen die Alternative $\theta=\theta_1$ zu vorgegebenem α und β bei festem Stichprobenumfang zu testen, ist der Stichprobenumfang so zu wählen, daß $V \geq (z_{1-\alpha} + z_{1-\beta})^2 / \theta_1^2$ ist (wobei für V der asymptotische Wert $I(0)$ zu nehmen ist). Die Nullhypothese ist abzulehnen, wenn bei positivem θ_1 $Z > z_{1-\alpha} \sqrt{V}$, bei negativem θ_1 $Z < -z_{1-\alpha} \sqrt{V}$ ist. $z_{1-\alpha}$ und $z_{1-\beta}$ sind die $1-\alpha$ bzw. $1-\beta$ Quantilen der Standard-Normalverteilung. Der Maximum-Likelihood-Schätzer für θ ist $Z/V (=l_{\theta}(0)/l_{\theta\theta}(0))$, ein $(1-\alpha)$ -Konfidenzintervall: $Z/V \pm z_{1-\alpha/2} / \sqrt{V}$.

Mit den Statistiken Z und V ist (unter Vernachlässigung von Ausdrücken der Ordnung $O(\theta^3)$) der Log-Likelihood-Quotient zum Test von $\theta=0$ gegen die Alternative $\theta=\theta_1$:

$$lr_m = \ln(LR_m) = \theta_1 Z_m - \frac{1}{2} \theta_1^2 V_m$$

wobei der Index m bei Z und V andeutet, daß diese Statistiken mit den Stichprobenwerten x_1, \dots, x_m zu bilden sind.

Der mit Z und V gebildete SPRT hat den Fortsetzungsbereich:

$$a_0 + bV_m < Z_m < a_1 + bV_m$$

mit:

$$a_0 = \frac{1}{\theta_1} \ln\left(\frac{\beta}{1-\alpha}\right); \quad a_1 = \frac{1}{\theta_1} \ln\left(\frac{1-\beta}{\alpha}\right); \quad b = \frac{1}{2} \theta_1$$

Der Bereich ist im Z - V -Koordinatensystem von parallelen Geraden begrenzt. Erreicht der Pfad die obere Gerade, so wird bei $\theta_1 > 0$ die Alternative angenommen, bei Erreichen der unteren Geraden die Nullhypothese. Bei $\theta_1 < 0$ ist bei Erreichen der unteren Geraden die Alternative anzunehmen und bei Erreichen der oberen Geraden die Nullhypothese.

Die Powerfunktion dieses sequentiellen Tests ist:

$$P(\theta) = \frac{1 - \left(\frac{\beta}{1-\alpha}\right)^{1-2\frac{\theta}{\theta_1}}}{\left(\frac{1-\beta}{\alpha}\right)^{1-2\frac{\theta}{\theta_1}} - \left(\frac{\beta}{1-\alpha}\right)^{1-2\frac{\theta}{\theta_1}}} \quad \text{für } \theta \neq 0.5\theta_1$$

und

$$P(\theta) = \frac{\ln\left(\frac{1-\alpha}{\beta}\right)}{\ln\left(\frac{1-\beta}{\alpha}\right) + \ln\left(\frac{1-\alpha}{\beta}\right)} \quad \text{für } \theta=0.5\theta_1$$

Die ASN-Funktion ist:

$$\text{ASN}(\theta) = \frac{\ln\left(\frac{1-\beta}{\alpha}\right) \left(1 - \left(\frac{\beta}{1-\alpha}\right)^{1-2\frac{\theta}{\theta_1}}\right) - \ln\left(\frac{\beta}{1-\alpha}\right) \left(1 - \left(\frac{1-\beta}{\alpha}\right)^{1-2\frac{\theta}{\theta_1}}\right)}{\theta_1(\theta - 0.5\theta_1) \left[\left(\frac{1-\beta}{\alpha}\right)^{1-2\frac{\theta}{\theta_1}} - \left(\frac{\beta}{1-\alpha}\right)^{1-2\frac{\theta}{\theta_1}}\right]} \quad \text{für } \theta \neq 0.5\theta_1$$

und

$$\text{ASN}(\theta) = \frac{1}{\theta_1^2} \ln\left(\frac{1-\alpha}{\beta}\right) \ln\left(\frac{1-\beta}{\alpha}\right) \quad \text{für } \theta=0.5\theta_1$$

Diese Formeln folgen aus den im Abschnitt 2 gebrachten allgemeinen Formeln für die Powerfunktion und ASN-Funktion des SPRT. Es wird dabei angenommen, daß zu einem gegebenen (ganzzahligen) V die Statistik Z als Summe von V unabhängigen und normal verteilten Zufallsgrößen U_i mit Mittelwert θ und Varianz 1 dargestellt werden kann; d. h. $Z_V = \sum u_i$, wobei u_i Realisation von U_i ist. Es gilt dann für den Log-Likelihood-Quotienten:

$$\text{lr}_V = \theta_1 Z_V - \frac{1}{2} \theta_1^2 V = \theta_1 (Z_V - \frac{1}{2} \theta_1 V) = \theta_1 \sum (u_i - \frac{1}{2} \theta_1) = \theta_1 \sum z_i.$$

Für die im Abschnitt 2 eingeführten Funktionen $h(\theta)$ und $E_\theta(z)$ ergibt dies: $h(\theta) = 1 - (2\theta/\theta_1)$ und $E(z_i|\theta) = \theta_1(\theta - \frac{1}{2}\theta_1)$. Damit ergeben sich die obigen Formeln für $P(\theta)$ und $\text{ASN}(\theta)$.

Erfolgt die Entscheidung über Fortsetzung oder Abschluß der Stichprobenerfassung nicht nach jedem Stichprobenergebnis, sondern jeweils bei Vorliegen einer Gruppe von Stichprobenwerten, dann kann ein 'overshooting' vorkommen; d. h. der Pfad ist am Ende der Gruppe bereits über die Entscheidungsgrenze hinausgelaufen. Man kann diesen Effekt durch eine Verengung des Fortsetzungsbereichs reduzieren, wobei die Reduktion dem bei der gegebenen Gruppengröße zu erwartenden Betrag des 'overshooting' entspricht. Unter der Annahme der Normalverteilung von Z ist dieser Erwartungswert $\delta a = 0.583\sqrt{\Delta V}$, wobei ΔV der (asymptotische) Wert der Information V für die Gruppe ist (Whitehead [7, 8]). Statt der

Achsenabschnitte a_0 und a_1 sind bei gruppensequentiellen Tests die Achsenabschnitte $a_0' = a_0 + \delta a$ bzw. $a_1' = a_1 - \delta a$ zu nehmen. Die Neigung b bleibt unverändert. Die Gruppengröße kann für jede Inspektion variabel sein. Der Wert ΔV ist dann für jede Gruppengröße getrennt zu berechnen.

Mit den Statistiken Z und V können für viele Testprobleme adäquate Sequenzverfahren konstruiert werden. Die Konstruktion von Z und V für normal verteilte und Bernoulli-verteilte Zufallsgrößen wird im folgenden Abschnitt 6 erläutert. Für Testvergleiche bei zwei Stichproben werden im Abschnitt 7 die Formeln für Z und V unter verschiedenen Verteilungsannahmen gebracht.

6. Z und V bei einer Stichprobe

6.1 Normal verteilte Größen

Die Stichprobenwerte x_1, \dots, x_m seien Realisationen von unabhängigen und identisch normal verteilten Zufallsgrößen mit Erwartungswert μ und Varianz σ^2 . Es wird zunächst angenommen, daß σ bekannt sei und die Nullhypothese $\mu=0$ gegen eine Alternative $\mu=\mu_1$ (>0) zu den Irrtumswahrscheinlichkeiten α und β getestet werden soll. Die log-Likelihood lautet:

$$l(\mathbf{x};\theta) = -\frac{1}{2}m \cdot \ln(2\pi\sigma^2) - \frac{1}{2\sigma^2} \sum_{i=1}^m (x_i - \mu)^2$$

wobei $\mathbf{x}=(x_1, \dots, x_m)'$ der Vektor der Stichprobenwerte ist.

Durch die Reparametrisierung: $\theta = \mu/\sigma$ erreicht man, daß bei vielen Testproblemen $|\theta| < 1$ angenommen werden kann. Mit dieser Reparametrisierung gilt:

$$l(\mathbf{x};\theta) = -\frac{1}{2}m \cdot \ln(2\pi\sigma^2) - \frac{1}{2} \sum_{i=1}^m \left(\frac{x_i}{\sigma} - \theta\right)^2$$

$$l_{\theta}(\mathbf{x};\theta) = \sum_{i=1}^m \left(\frac{x_i}{\sigma} - \theta\right) = \frac{1}{\sigma} \sum_{i=1}^m x_i - m \cdot \theta$$

$$l_{\theta\theta}(\mathbf{x};\theta) = -m$$

Daraus folgt:

$$Z = l_{\theta}(\mathbf{x};0) = \frac{1}{\sigma} \sum_{i=1}^m x_i \quad V = m$$

Ist σ unbekannt, dann liegt ein Störparameter vor. Mit der Reparametrisierung $\theta=\mu/\sigma$ und $\psi=1/\sigma$ lautet die log-Likelihood:

$$l(\mathbf{x};\theta,\psi) = -\frac{1}{2}m \cdot \ln(2\pi) + m \cdot \ln(\psi) - \frac{1}{2} \sum_{i=1}^m (\psi x_i - \theta)^2$$

Damit ist:

$$l_{\theta}(\mathbf{x};\theta,\psi) = \psi \sum_{i=1}^m x_i - m\theta$$

$$l_{\psi}(\mathbf{x};\theta,\psi) = \frac{m}{\psi} - \psi \sum_{i=1}^m x_i^2 + \theta \sum_{i=1}^m x_i$$

$$l_{\theta\theta}(\mathbf{x};\theta,\psi) = -m$$

$$l_{\theta\psi}(\mathbf{x};\theta,\psi) = \sum_{i=1}^m x_i$$

$$l_{\psi\psi}(\mathbf{x};\theta,\psi) = -\frac{\psi^2 \sum_{i=1}^m x_i^2 + m}{\psi^2}$$

ML-Schätzer für ψ bei gegebenem θ ist Lösung der quadratischen Gleichung:

$$\psi^2 \sum_{i=1}^m x_i^2 - \psi \theta \sum_{i=1}^m x_i - m = 0$$

Die positive Lösung dieser Gleichung ist:

$$\hat{\psi}(\theta) = \frac{1}{2} \theta \frac{\sum_{i=1}^m x_i}{\sum_{i=1}^m x_i^2} + \frac{1}{2 \sum_{i=1}^m x_i^2} \sqrt{\theta^2 \left(\sum_{i=1}^m x_i \right)^2 + 4m \sum_{i=1}^m x_i^2}$$

Daraus folgt:

$$\psi^* = \hat{\psi}(0) = \sqrt{\frac{m}{\sum_{i=1}^m x_i^2}} = \frac{1}{D} \quad \text{mit } D = \sqrt{\frac{1}{m} \sum_{i=1}^m x_i^2}$$

Es ist somit:

$$Z = l_{\theta}(\mathbf{x}; \psi^*) = \frac{\sum_{i=1}^m x_i}{D}$$

$$V = - (l_{\theta\theta}(\mathbf{x}; 0, \psi^*) - l_{\theta\psi}(\mathbf{x}; 0, \psi^*)^2 / l_{\psi\psi}(\mathbf{x}; 0, \psi^*)) = m - \frac{\left(\sum_{i=1}^m x_i \right)^2}{2 \sum_{i=1}^m x_i^2} = m - \frac{Z^2}{2m}$$

Für $|\theta| \ll 1$ bleibt die Statistik Z dem Betrage nach klein, so daß in diesem Fall für $m \rightarrow \infty$ $V \rightarrow m$ gilt.

6.2 Bivariate (Bernoulli-verteilte) Größen

Die Stichprobenwerte x_1, \dots, x_m sollen Realisationen unabhängiger und identisch Bernoulli-verteilter Zufallsgrößen sein; d.h. sie nehmen nur die Werte 0 und 1 an mit $\Pr(X=1)=\pi$. Zu testen ist die Nullhypothese $\pi=\pi_0$ gegen eine Alternative $\pi=\pi_1 (>\pi_0)$. Wurde in der Stichprobe r -mal die 1 und $(n-r)$ -mal die 0 beobachtet, dann lautet die log-Likelihood:

$$l(\mathbf{x}; \pi) = \text{const} + r \cdot \ln(\pi/(1-\pi)) + m \cdot \ln(1-\pi).$$

Nach der Reparametrisierung:

$$\theta = \ln \frac{\pi(1-\pi_0)}{\pi_0(1-\pi)}$$

lautet die log-Likelihood:

$$l(\mathbf{x}; \theta) = \text{const} + r(\theta + \gamma) - m \cdot \ln(1 + e^{\theta + \gamma}) \quad \text{mit } \gamma = \ln(\pi_0/(1-\pi_0))$$

Es folgt:

$$l_{\theta}(\mathbf{x}; \theta) = r - m \frac{e^{\theta + \gamma}}{1 + e^{\theta + \gamma}} \quad \text{und} \quad l_{\theta\theta}(\mathbf{x}; \theta) = -m \frac{e^{\theta + \gamma}}{(1 + e^{\theta + \gamma})^2}$$

Damit ist:

$$Z = I_{\theta}(\mathbf{x};0) = r - m \frac{e^{\gamma}}{1 + e^{\gamma}} = r - m\pi_0$$

$$V = -I_{\theta\theta}(\mathbf{x};0) = m \frac{e^{\gamma}}{(1 + e^{\gamma})^2} = m\pi_0(1-\pi_0)$$

7 Die Statistiken Z und V bei 2-Stichproben-Vergleichen

Bei einem 2-Stichproben-Vergleich liegen n_1 Meßwerte x_{1i} der ersten Stichprobe und n_2 Meßwerte x_{2i} der zweiten Stichprobe vor. Die Meßwerte der ersten Stichprobe sind unabhängige Realisationen mit der Verteilungsdichte $f(x;\delta_1,\psi)$; die der zweiten Stichprobe haben die Verteilungsdichte $f(y;\delta_2,\psi)$. δ_1 und δ_2 sind die interessierenden Parameter und ψ soll Störparameter symbolisieren. Zu testen ist die Nullhypothese $H_0: \delta_1=\delta_2$ gegen eine (einseitige) Alternative, z. B. $\delta_1>\delta_2$. Um zu einer einfachen Nullhypothese zu kommen, ist eine Reparametrisierung erforderlich; d. h. es wird ein neuer Parameter θ eingeführt, so daß der zusammengesetzten Nullhypothese H_0 die einfache Nullhypothese $H_0': \theta=0$ entspricht. Als Alternative wird ein Wert θ_1 fixiert, der durch Vorgaben für δ_1 und δ_2 berechnet werden kann. Im allgemeinen wird dabei ein neuer Störparameter ϕ eingeführt. Empfehlenswert ist folgendes Vorgehen:

Die Parameter θ und ϕ werden definiert als: $\theta=1/2(\delta_1-\delta_2)$ und $\phi=1/2(\delta_1+\delta_2)$. Für die ursprünglichen Parameter δ_1 und δ_2 gilt dann: $\delta_1=\phi+\theta$ und $\delta_2=\phi-\theta$. Für die Log-Likelihood-Funktion gilt:

$$l(\mathbf{x},\mathbf{y};\theta,\phi,\psi) = l^{(x)}(\mathbf{x};\delta_1,\psi) + l^{(y)}(\mathbf{y};\delta_2,\psi)$$

Daraus folgt für die Ableitungen (wobei die Argumente und Parameter weggelassen werden):

$$\begin{aligned} I_{\theta} &= I_{\delta_1}^{(x)} - I_{\delta_2}^{(y)} \\ I_{\phi} &= I_{\delta_1}^{(x)} + I_{\delta_2}^{(y)} \\ I_{\psi} &= I_{\psi}^{(x)} + I_{\psi}^{(y)} \\ I_{\theta\theta} &= I_{\phi\phi} = I_{\delta_1\delta_1}^{(x)} + I_{\delta_2\delta_2}^{(y)} \\ I_{\theta\phi} &= I_{\delta_1\delta_1}^{(x)} - I_{\delta_2\delta_2}^{(y)} \\ I_{\theta\psi} &= I_{\delta_1\psi}^{(x)} - I_{\delta_2\psi}^{(y)} \\ I_{\phi\psi} &= I_{\delta_1\psi}^{(x)} + I_{\delta_2\psi}^{(y)} \\ I_{\psi\psi} &= I_{\psi\psi}^{(x)} + I_{\psi\psi}^{(y)} \end{aligned}$$

Die Schätzwerte ϕ^* und ψ^* sind Lösungen der beiden Gleichungen:

$$I_{\delta_1}^{(x)}(\phi^*,\psi^*) + I_{\delta_2}^{(y)}(\phi^*,\psi^*) = 0 \quad \text{und} \quad I_{\psi}^{(x)}(\phi^*,\psi^*) + I_{\psi}^{(y)}(\phi^*,\psi^*) = 0$$

Damit können die Statistiken Z und V berechnet werden.

Im folgenden werden Formeln für Z und V für einige 2-Stichproben-Vergleiche angegeben.

7.1 Binäre Zielgrößen

In der ersten Stichprobe wurden bei n_1 Beobachtungen r_1 Erfolge festgestellt, in der zweiten Stichprobe bei n_2 Beobachtungen r_2 Erfolge. Die Wahrscheinlichkeit für einen Erfolg ist in der ersten Stichprobe π_1 und in der zweiten Stichprobe π_2 . Zu testen ist die Nullhypothese $\pi_1 = \pi_2$ gegen eine einfache Alternative.

Testparameter:

$$\theta = \ln \left(\frac{\pi_1(1 - \pi_2)}{\pi_2(1 - \pi_1)} \right)$$

Damit ist:

$$Z = \frac{n_2 r_1 - n_1 r_2}{n_1 + n_2}$$

und

$$V = \frac{n_1 n_2 (r_1 + r_2)(n_1 + n_2 - (r_1 + r_2))}{(n_1 + n_2)^3}$$

Werden n_1 und n_2 im festen Verhältnis $R (= n_1/n_2)$ vergrößert, so gilt (mit $N = n_1 + n_2$) asymptotisch für $N \rightarrow \infty$:

$$V \rightarrow N \frac{R}{(1+R)^2} \pi_1(1 - \pi_1) \quad \text{für } \pi_1 = \pi_2$$

$$V \rightarrow N \frac{R}{(1+R)^4} (R\pi_1 + \pi_2)(1 + R - (R\pi_1 + \pi_2)) \quad \text{für } \pi_1 \neq \pi_2$$

7.2 Quantitative (normal verteilte) Zielgrößen

Es liegen 2 unabhängige Stichproben quantitativer Größen: $\{x_{1i}\}$ und $\{x_{2i}\}$ vom Umfang n_1 bzw. n_2 vor. Die x_{1i} sind unabhängige Realisationen normal verteilter Zufallsgrößen mit Mittelwert μ_1 und Varianz σ^2 , die x_{2i} von normal verteilten Zufallsgrößen mit Mittelwert μ_2 und Varianz σ^2 . Zu testen ist die Hypothese $H_0: \mu_1 = \mu_2$ gegen eine einfache Alternative.

Testparameter ist:

$$\theta = \frac{\mu_1 - \mu_2}{\sigma}$$

Man berechne die Mittelwerte \bar{x}_1 und \bar{x}_2 der beiden Stichproben und den M-L-Schätzer für σ^2 bei $\theta=0$:

$$S^2 = (\sum x_{1i}^2 + \sum x_{2i}^2 - (\sum x_{1i} + \sum x_{2i})^2 / (n_1 + n_2)) / (n_1 + n_2)$$

Damit ist:

$$Z = \frac{n_1 n_2}{n_1 + n_2} \cdot \frac{\bar{x}_1 - \bar{x}_2}{S}$$

$$V = \frac{n_1 n_2}{n_1 + n_2} - \frac{Z^2}{2(n_1 + n_2)}$$

Werden n_1 und n_2 im festen Verhältnis $R(=n_1/n_2)$ vergrößert, so gilt (mit $N=n_1+n_2$) asymptotisch:

$$V_{\text{asympt}} \approx \frac{R}{(R+1)^2} \cdot N \quad (= \frac{N}{4} \text{ für } n_1 = n_2 = N/2)$$

7.3 Ordinale Zielgrößen

Es liegen 2 unabhängige Stichproben $\{x_{1i}\}$ und $\{x_{2i}\}$ vom Umfang n_1 bzw. n_2 vor. Die Stichprobenwerte seien Realisationen von unabhängigen Zufallsgrößen X_1 bzw. X_2 , die über $k (>2)$ ordinale Kategorien C_1, C_2, \dots, C_k diskret verteilt sind, mit den Wahrscheinlichkeiten: $P(X_1=C_r)=\pi_{1r}$ und $P(X_2=C_r)=\pi_{2r}$. Es werden die kumulativen Wahrscheinlichkeiten für Kategorien kleiner oder gleich C_j mit $Q_{1j}=\pi_{11}+\pi_{12}+\dots+\pi_{1j}$ und $Q_{2j}=\pi_{21}+\pi_{22}+\dots+\pi_{2j}$ ($j=1, \dots, k-1$) bezeichnet.

Testparameter ist:

$$\theta = \ln \frac{Q_{1j}(1-Q_{2j})}{Q_{2j}(1-Q_{1j})}$$

Es wird angenommen, daß dieser Parameter für alle j denselben Wert hat, so daß ein beliebiges j ($=1, \dots, k-1$) gewählt werden kann (proportional odds ratio).

Bezeichnet n_{1r} die Zahl der Beobachtungen der ersten Stichprobe, die in die Kategorie C_r fallen, und n_{2r} die der zweiten Stichprobe mit dieser Eigenschaft ($r=1, \dots, k$; $\sum n_{1r}=n_1$, $\sum n_{2r}=n_2$), dann werden folgende Größen berechnet:

$$\begin{aligned} L_{1j} &= n_{11} + \dots + n_{1(j-1)}; & L_{2j} &= n_{21} + \dots + n_{2(j-1)}; & \text{für } j=2, \dots, k & & L_{11} &= L_{21} &= 0 \\ U_{1j} &= n_{1(j+1)} + \dots + n_{1k}; & U_{2j} &= n_{2(j+1)} + \dots + n_{2k}; & \text{für } j=1, \dots, k-1; & & U_{11} &= U_{21} &= 0 \\ N_j &= n_{1j} + n_{2j} & \text{für } j=1, \dots, k; & & N &= \sum N_j \end{aligned}$$

Damit lauten die Formeln für die Statistiken Z und V :

$$Z = \frac{1}{N+1} \sum_{j=1}^k n_{2j} (L_{1j} - U_{1j}) = - \frac{1}{N+1} \sum_{j=1}^k n_{1j} (L_{2j} - U_{2j})$$

$$V = \frac{Z^2}{N+2} + \frac{W}{(N+1)(N+2)}$$

mit

$$W = \sum_{j=1}^k (n_{1j}(n_2 - n_{2j}) + n_{1j}n_{2j}(N - N_j) + 2n_{1j}L_{2j}U_{2j} + 2n_{2j}L_{1j}U_{1j})$$

Man beachte, daß Z der Mann-Whitney-U-Statistik entspricht.

Werden n_1 und n_2 im festen Verhältnis $R(=n_1/n_2)$ vergrößert, so gilt asymptotisch unter H_0 :

$$V_{asympt} \approx \frac{RN}{3(R+1)^2} (1 - \sum_{j=1}^k \pi_j^3) \quad (= \frac{N}{12} (1 - \sum_{j=1}^k \pi_j^3) \text{ für } n_1 = n_2 = N/2)$$

wobei π_j die Response-Wahrscheinlichkeiten für die Kategorien C_j unter H_0 sind. Diese Werte sind problemadäquat vorzugeben.

Der Test kann auch als „parameterfreier“ Test für Unterschiede in den Verteilungen zweier quantitativer Stichproben benutzt werden. Die k Kategorien werden dann von den verschiedenen Stichprobenwerten beider Stichproben gebildet.

7.4 Zensierte Größen (Überlebenszeiten)

Als Beispiel wird angenommen, daß eine Patientenstichprobe vom Umfang n_1 mit einem Arzneimittel A und eine zweite davon unabhängige Stichprobe vom Umfang n_2 mit dem Mittel B behandelt wurde. Beobachtet werden die Überlebenszeiten t_i , die auch zensiert sein können; d. h. bei einem Teil der Patienten ist nur die Zeit der letzten Visite bekannt und zu dieser Zeit haben sie noch gelebt. Die Überlebenszeiten der ersten Stichproben sind unabhängige Realisationen einer Zufallsgröße T_1 mit der Verteilungsfunktion $F_1(t)$, die der zweiten von T_2 mit der Verteilungsfunktion $F_2(t)$ ($t \geq 0$). Es wird das proportional-hazard-rate-Modell ($\lambda_1(t) = \lambda_2(t) \exp(\theta)$), wobei $\lambda_i(t)$ die Hazardfunktion der Gruppe i ($i=1,2$) ist) angenommen. Damit ist:

$$1 - F_1(t) = (1 - F_2(t))^{\exp(\theta)}$$

Der Parameter θ ist die log-hazard-rate. Für $\theta=0$ sind beide Verteilungen gleich, für $\theta>0$ sind die Überlebenszeiten in Gruppe 1 kürzer ($F_1(t) > F_2(t)$) und für $\theta<0$ länger ($F_1(t) < F_2(t)$). Zu testen ist die Nullhypothese $\theta=0$ gegen eine einfache Alternative θ_1 .

Der Parameter θ_1 kann mit den beiden Überlebensraten $Q_1(t_0)$ und $Q_2(t_0)$ für eine willkürliche Beobachtungsperiode t_0 festgelegt werden. Es ist: $Q_1(t_0) = 1 - F_1(t_0)$ und $Q_2(t_0) = 1 - F_2(t_0)$. Daraus folgt:

$$\theta = \ln \frac{\ln(Q_1(t_0))}{\ln(Q_2(t_0))} \quad \text{unabhängig von } t_0$$

Es wurden in beiden Stichproben insgesamt k verschiedene, nicht-zensierte Überlebenszeiten beobachtet. Diese werden der Größe nach geordnet: $t_1 < t_2 < \dots < t_k$. Die folgenden Bezeichnungen werden eingeführt:

- e_{1i} = Zahl der Patienten der Gruppe 1, die unmittelbar vor t_i noch leben
- e_{2i} = Zahl der Patienten der Gruppe 2, die unmittelbar vor t_i noch leben

$$\begin{aligned}
d_{1i} &= \text{Zahl der Patienten der Gruppe 1, die zur Zeit } t_i \text{ sterben} \\
d_{2i} &= \text{Zahl der Patienten der Gruppe 2, die zur Zeit } t_i \text{ sterben} \\
E_i &= e_{1i} + e_{2i} \\
D_i &= d_{1i} + d_{2i}
\end{aligned}$$

Mit diesen Größen lauten die Formeln für Z und V:

$$\begin{aligned}
Z &= \sum_{i=1}^k \frac{e_{2i} d_{1i} - e_{1i} d_{2i}}{E_i} \\
V &= \sum_{i=1}^k \frac{D_i (E_i - D_i) e_{1i} e_{2i}}{E_i^2 (E_i - 1)}
\end{aligned}$$

Die Zahl e_{1i} bzw. e_{2i} gibt die Zahl der Patienten an, die zur Zeit t_i in der Gruppe 1 bzw. 2 dem Sterberisiko ausgesetzt sind. Das sind alle Patienten der jeweiligen Gruppe, deren Sterbezeit oder Zensierungszeit (Zeit der letzten Beobachtung) größer oder gleich t_i ist.

Wenn zu den Zeiten t_i jeweils nur ein Patient der Gruppe 1 oder 2 stirbt (d. h. d_{1i} und d_{2i} sind entweder 1 oder 0 und $D_i=1$ für alle i) dann gilt für V:

$$V \approx \sum_{i=1}^k \frac{e_{1i} e_{2i}}{E_i^2}$$

Wenn für alle Sterbezeiten t_i das Verhältnis e_{1i}/e_{2i} der 'exponierten' Patienten annähernd mit der Allokationsrate $R=n_1/n_2$ übereinstimmt, dann folgt aus dieser Formel:

$$V_{\text{asympt}} \approx \frac{R}{(R+1)^2} ED(t_0)$$

$ED(t_0)$ ist der Erwartungswert der Anzahl k aller Ereignisse (Tod) in der Beobachtungszeit t_0 . Dieser Erwartungswert ist:

$$ED(t_0) = n_1(1 - Q_1(t_0)) + n_2(1 - Q_2(t_0)) = \frac{N(t_0)}{R+1} (R(1 - Q_1(t_0)) + 1 - Q_2(t_0))$$

wobei $Q_1(t_0)$ und $Q_2(t_0)$ die Überlebensraten für die Gruppe 1 und 2 in einer willkürlichen Beobachtungszeit t_0 und $N(t_0)$ die Gesamtzahl der über die Zeit t_0 beobachteten Patienten sind.

Daraus ergibt sich asymptotisch für V:

$$V_{\text{asympt}} \approx \frac{RN(t_0)}{(1+R)^3} (R(1 - Q_1(t_0)) + 1 - Q_2(t_0)) \quad (= \frac{N(t_0)}{8} (2 - Q_1(t_0) - Q_2(t_0)) \text{ für } n_1=n_2=N)$$

8 Sequentielle Dreieckspläne (geschlossene Pläne)

Ein wesentlicher Nachteil des SPRT besteht darin, daß er eine offene Fortsetzungsregion besitzt. Wenn auch der Sequenzpfad mit Wahrscheinlichkeit 1 nach endlich vielen Schritten zu einer Entscheidung führt, so kann dies aber sehr lange dauern. Wald [6] hat deshalb

vorgeschlagen, die Sequenzprozedur nach einer endlichen Zahl n^* zu beenden und für H_0 zu entscheiden, wenn $\ln B < \ln r_{n^*} \leq 0$ ist, oder für die Alternative zu entscheiden, wenn $0 < \ln r_{n^*} < \ln A$ ist. Damit gelten für die Irrtumswahrscheinlichkeiten nicht mehr die vorgegebenen Grenzen α und β . Diese können selbst bei $n^*=1000$ noch beträchtlich (bis zum Doppelten) überschritten werden. Armitage [2] hat vorgeschlagen, die Begrenzungsgeraden des Fortsetzungsbereichs so zu verändern, daß bei einer Beendigung des Verfahrens nach n^* Schritten (es soll dann die Nullhypothese angenommen werden, wenn vorher keine der Grenzen erreicht wurde) die Irrtumswahrscheinlichkeiten α und β eingehalten werden. Als Zusatzbedingung hat er dabei festgelegt, daß über den gesamten Verlauf der oberen Begrenzung der Likelihoodquotient konstant gleich $(1-\beta)/\alpha$ sein soll. Er hat approximativ die Lage der Begrenzungsgeraden bestimmt, die dieser Bedingung genügen und α und β einhalten.

Ein anderer Vorschlag für geschlossene Sequenzpläne wurde vor allem von Whitehead [7, 8] unter Verwendung der Statistiken Z und V propagiert. Bei diesen Plänen wird der Fortsetzungsbereich durch ein Dreieck begrenzt, wie dies in Abb. 2 gezeigt ist. Die Pläne heißen deshalb **Dreieckspläne**. Diese Pläne sind primär für symmetrische Irrtumswahrscheinlichkeiten $\alpha=\beta$ entwickelt. Das Dreieck, das den Fortsetzungsbereich begrenzt, hat eine Seite auf der Z -Achse von $-a$ bis $+a$. Die beiden anderen Seiten werden durch eine Größe c bestimmt (vergl. Abb.2).

Die Entscheidungsregel zur Entscheidung von $H_0: \theta=0$ gegen $H_1: \theta=\theta_1$ lautet: Setze die Beobachtung fort, wenn:

$$-a+3cV_m < Z_m < a+cV_m$$

Für $\theta_1 > 0$ ist H_0 anzunehmen, wenn $Z_m \leq -a+3cV_m$, und H_1 anzunehmen, wenn $Z_m \geq a+cV_m$; für $\theta_1 < 0$ ist umgekehrt zu verfahren.

Die Irrtumswahrscheinlichkeiten α und $\beta=\alpha$ werden eingehalten, wenn für a und c gilt:

$$a = 2\ln(1/2\alpha)/\theta_1, \quad c = \theta_1/4$$

Die beiden Geraden, die den Fortsetzungsbereich begrenzen, haben die Achsenabschnitte $-a$ und a und schneiden sich im Punkt (Z_{\max}, V_{\max}) mit

$$V_{\max} = a/c = 8\ln(1/2\alpha)/\theta_1^2; \quad Z_{\max} = 2a = 4\ln(1/2\alpha)/\theta_1$$

Wird dieser Schnittpunkt exakt erreicht, dann ist die Alternative anzunehmen.

Als Beispiel sei ein Vergleich der Wahrscheinlichkeiten π_1 und π_2 zweier binärer Variablen betrachtet. Zu testen ist die Nullhypothese: $\pi_1=\pi_2$ gegen die Alternative: $\pi_1=0.6$ und $\pi_2=0.5$. Testparameter ist die log-odds-ratio (vergl. Abschnitt 6.2) mit dem Referenzwert $\theta_1=\ln(0.6 \cdot 0.5 / (0.4 \cdot 0.5))=0.405$. Damit folgt für $\alpha=\beta=0.05$:

$$a = 11.36, \quad c = 0.101, \quad V_{\max}=112.48, \quad Z_{\max}=22.72$$

Für $\pi_0=(\pi_1+\pi_2)/2=0.55$ entspricht dem ein $N_{\max}=4V_{\max}/(\pi_0(1-\pi_0))=1818$ (bei $n_1=n_2$). Bei festem Stichprobenumfang ist ein $V_{\text{fix}}=(z_{1-\alpha}+z_{1-\beta})^2/\theta_1^2=65.5$ oder $N_{\text{fix}}=1060$ erforderlich. Der Sequentialplan kommt im Mittel bei $V_{\text{exp}}=35.18$ zu einer Entscheidung, wenn die Null- oder Alternativhypothese zutrifft. Dies entspricht einem Gesamtstichprobenumfang von $N_{\text{exp}}=569$. Falls $\theta=1/2\theta_1$ ist, endet das Verfahren im Mittel bei $V_{\text{exp}}=47.06$ bzw. $N_{\text{exp}}=760$ Beobachtungen. Im Mittel kann also beim Sequenzverfahren fast die Hälfte der für feste Stichproben erforderlichen Zahl von Beobachtungen eingespart werden. Maximal kann allerdings beim Dreiecksplan der Stichprobenumfang um ca. 80% größer werden als bei festem Stichprobenumfang. Dies ist allerdings sehr unwahrscheinlich.

Verfahren mit unsymmetrischen Irrtumswahrscheinlichkeiten ($\alpha \neq \beta$) können durch eine Modifikation des Referenzwertes θ_1 aus den symmetrischen Verfahren abgeleitet werden. Entsprechend einem Vorschlag bei Whitehead [8] ist ein symmetrischer Dreiecksplan mit dem gegebenen α und einem modifizierten θ_1' zu konstruieren, der für diesen Referenzwert und $\beta'=\alpha$ dasselbe V_{fix} ergibt, wie es bei festem Stichprobenumfang für α , β und θ_1 erforderlich ist. Der modifizierte Referenzwert θ_1' wird somit aus der Gleichung:

$$(2z_{1-\alpha}/\theta_1')^2 = (z_{1-\alpha}+z_{1-\beta})^2/\theta_1^2$$

bestimmt. Dies ergibt:

$$\theta_1' = 2\theta_1/(1+z_{1-\beta}/z_{1-\alpha}),$$

wobei $z_{1-\alpha}$ und $z_{1-\beta}$ die $1-\alpha$ - bzw. $1-\beta$ -Quantilen der Standard-Normalverteilung sind. Die Kenngrößen a und c der Begrenzungsgeraden sind somit bei unsymmetrischem α und β :

$$a = (1+z_{1-\beta}/z_{1-\alpha})\ln(1/2\alpha)/\theta_1, \quad c = \theta_1/(2(1+z_{1-\beta}/z_{1-\alpha}))$$

Wählt man im obigen Beispiel $\alpha=0.05$ und $\beta=0.20$, so ergibt sich ein modifiziertes $\theta_1'=0.54$ und damit $a=8.58$, $c=0.13$, $N_{\max}=1035$, $N_{\text{fix}}=608$, $N_{\text{exp}}=325$ für $\theta=0$ oder $\theta=\theta_1$ und $N_{\text{exp}}=434$ für $\theta=1/2\theta_1$.

Nimmt man die standardisierten Koordinaten $Z_s = \theta_1 Z$ und $V_s = \theta_1^2 V$, dann erhält man Begrenzungsgeraden, die unabhängig von θ_1 sind und nur noch von α und β abhängen. Die Kenngrößen der standardisierten Begrenzungsgeraden sind:

$$as = (1+z_{1-\beta}/z_{1-\alpha})\ln(1/2\alpha) \text{ und } cs = 1/(2(1+z_{1-\beta}/z_{1-\alpha}))$$

Die Begrenzungsgeraden des Fortsetzungsbereichs im standardisierten (Z_s, V_s) -Koordinatensystem sind damit:

$$\text{obere Begrenzung: } Z_s = as + cs \cdot V_s, \quad \text{untere Begrenzung: } Z_s = -as + 3cs \cdot V_s$$

Es soll nun gezeigt werden, daß ein symmetrischer Dreieckstest ($\alpha=\beta$) die Bedingungen: $P(0) \leq \alpha$ und $OC(\theta_1) \leq \alpha$ erfüllt. Es bedeutet $P(\theta)$ die Powerfunktion (Wahrscheinlichkeit H_1 anzunehmen, wenn θ gilt) und $OC(\theta)$ die OC-Funktion (Wahrscheinlichkeit H_0 anzunehmen, wenn θ gilt). Als Fortsetzungsbereich wird der Bereich: $-a+3/4\theta_1 V_m < Z_m < a+1/4\theta_1 V_m$ mit unbekanntem a angesetzt. Für das symmetrische Verfahren soll die Symmetrieforderung: $P(\theta_1-\theta)=1-P(\theta)$ und $OC(\theta_1-\theta)=1-OC(\theta)$ gelten. Insbesondere gilt: $OC(1/2\theta_1)=P(1/2\theta_1)=1/2$.

Für die Herleitung wird angenommen, daß die Statistik V_m durch ihren Erwartungswert mI (mit $I=I(0)$ =Fisher Information von x für $\theta=0$) ersetzt werden kann und für $f(\mathbf{x}_m; \theta)$ die Approximation: $f(\mathbf{x}_m; \theta) \sim \exp(\theta Z_m - \frac{1}{2} \theta^2 mI)$ gilt (\sim bedeutet 'proportional').

Die Hypothese H_0 wird nach m Beobachtungen angenommen, wenn $Z_m \leq -a + \frac{3}{4} \theta_1 mI$ ist. Die Menge der Punkte $\mathbf{x}_m = (x_1, \dots, x_m)'$, für die H_0 angenommen wird, wird mit B_m bezeichnet. Es ist:

$$OC(\theta_1) = \Pr(\text{Annahme } H_0 | \theta_1) = \sum_{m=1}^{\infty} \int_{B_m} f(\mathbf{x}_m; \theta_1) d\mathbf{x}_m$$

Aus $f(\mathbf{x}_m; \theta) \sim \exp(\theta Z_m - \frac{1}{2} \theta^2 mI)$ folgt: $f(\mathbf{x}_m; \theta_1) / f(\mathbf{x}_m; \frac{1}{2} \theta_1) = \exp(\frac{1}{2} \theta_1 Z_m - (3/8) \theta_1^2 mI)$.

Für alle Punkte \mathbf{x}_m aus B_m ist (wegen $Z_m \leq -a + \frac{3}{4} \theta_1 mI$): $f(\mathbf{x}_m; \theta_1) \leq \exp(-\frac{1}{2} \theta_1 a) f(\mathbf{x}_m; \frac{1}{2} \theta_1)$ und:

$$OC(\theta_1) \leq \exp(-\frac{1}{2} \theta_1 a) \sum_{m=1}^{\infty} \int_{B_m} f(\mathbf{x}_m; \frac{1}{2} \theta_1) d\mathbf{x}_m = \exp(-\frac{1}{2} \theta_1 a) OC(\frac{1}{2} \theta_1) = \frac{1}{2} \exp(-\frac{1}{2} \theta_1 a)$$

Da $OC(\theta_1) \leq \alpha$ sein soll, folgt daraus: $a = (2/\theta_1) \ln(1/2\alpha)$.

Die Alternative H_1 wird bei einem gegebenen m für die Menge A_m von \mathbf{x}_m -Werten angenommen, für die $Z_m \geq a + \frac{1}{4} \theta_1 mI$ gilt. Es ist somit:

$$P(0) = \Pr(\text{Annahme } H_1 | 0) = \sum_{m=1}^{\infty} \int_{A_m} f(\mathbf{x}_m; 0) d\mathbf{x}_m$$

Für alle \mathbf{x}_m aus A_m (mit $Z_m \geq a + \frac{1}{4} \theta_1 mI$) ist: $f(\mathbf{x}_m; 0) \leq \exp(-\frac{1}{2} \theta_1 a) f(\mathbf{x}_m; \frac{1}{2} \theta_1)$ und damit:

$$P(0) \leq \exp(-\frac{1}{2} \theta_1 a) \sum_{m=1}^{\infty} \int_{A_m} f(\mathbf{x}_m; \frac{1}{2} \theta_1) d\mathbf{x}_m = \exp(-\frac{1}{2} \theta_1 a) P(\frac{1}{2} \theta_1) = \frac{1}{2} \exp(-\frac{1}{2} \theta_1 a)$$

Mit $a = (2/\theta_1) \ln(1/2\alpha)$ ist somit auch die Bedingung: $P(0) \leq \alpha$ erfüllt.

Für die Verteilung $F(v^*; \theta)$ der Zufallsgröße $V^* = m^* I$, bei der der Sequenzpfad eine der Entscheidungsgrenzen erreicht, hat Whitehead in [8] Reihenentwicklungen hergeleitet. $F(v^*; \theta)$ ist die Wahrscheinlichkeit, daß ein Pfad bei einem $V^* \leq v^*$ eine der Grenzen erreicht, wenn der Parameterwert θ ist. Der Pfad kann nach Erreichen der Grenze im betreffenden Annahmehbereich verbleiben, in den Fortsetzungsbereich zurückkehren, erneut die betreffende Annahmehgrenze kreuzen oder die andere Annahmehgrenze kreuzen; und das ganze kann sich beliebig häufig wiederholen. Die Wahrscheinlichkeiten für diese verschiedenen Möglichkeiten bilden die Glieder der Reihe.

Die Reihe wurde von Whitehead hergeleitet. Die Glieder der Reihe hängen von den Kenngrößen des Plans: $a = (1 + z_{1-\beta} / z_{1-\alpha}) \ln(1/2\alpha) / \theta_1$ und $c = \theta_1 / 4$ ab. Die Verteilungsfunktion $F(v; \theta)$ von V^* ist die Summe aus der Wahrscheinlichkeit $Q(0, v; \theta)$, H_0 bei einem $V^* \leq v$ anzunehmen, wenn der Parameter den Wert θ hat, und der Wahrscheinlichkeit $Q(1, v; \theta)$, H_1 bei einem $V^* \leq v$ anzunehmen, wenn der Parameter den Wert θ hat:

$F(v; \theta) = Q(0, v; \theta) + Q(1, v; \theta)$. Für die Wahrscheinlichkeit $Q(0, v; \theta)$ gilt folgende Reihenentwicklung:

$$Q(0, v; \theta) = \varphi \left(\frac{(\theta - 3c)v + a}{\sqrt{v}} \right) \sum_{s=0}^{\infty} (-1)^s \exp(2s(s+1)a(c - a/v)) \left\{ R \left(\frac{(\theta - 3c)v + (2s+1)a}{\sqrt{v}} \right) + R \left(\frac{-(\theta - 3c)v + (2s+1)a}{\sqrt{v}} \right) \right\}$$

Die Funktion $\varphi(\cdot)$ ist die Dichte der Standard-Normalverteilung und $R=(1-\Phi(\cdot))/\varphi(\cdot)$.
 Aus Symmetriegründen gilt: $Q(1,v;\theta)=\Pr(V^*\leq v;|\text{Entscheidung } H_1;\theta)=Q(0,v;4c-\theta)$. Daraus folgt für die Verteilungsfunktion $F(v;\theta)$ von V^* :

$$F(v;\theta) = Q(0,v;\theta)+Q(0,v;4c-\theta)$$

Aus der Reihe für $F(v^*;\theta)$ kann die Momente-erzeugende-Funktion (moment generating function) und daraus die ASN-Funktion $E(V^*|\theta)$ hergeleitet werden.

Die Formel für die ASN-Funktion lautet:

$$\text{ASN}(\theta) = E(V^*|\theta) = e(\theta) + e(\theta_1-\theta)$$

mit:

$$e(\theta) = \frac{a\varphi((\theta - 2c)\sqrt{a/c})}{3c - \theta} \sum_{s=1}^{\infty} (-1)^s (2s+1) \left\{ R((\theta - 2c)\sqrt{a/c} + 2s\sqrt{ac}) - R(-(\theta - 2c)\sqrt{a/c} + 2(s+1)\sqrt{ac}) \right\}$$

für $\theta \neq 3c$ und

$$e(3c) = 2a\varphi(\sqrt{ac}) \sum_{s=1}^{\infty} (-1)^s \left\{ (2s+1)\sqrt{a/c} - (2s+1)^2 aR((2s+1)\sqrt{ac}) \right\}$$

Der Dreiecksplan besitzt (für $\alpha \leq 0.05$) fast minimale ASN für $\theta = 1/2\theta_1$ (im Vergleich zu anderen Sequenzplänen) (vergl. [8]).

Erfolgt die sequentielle Testung in Gruppen, bei denen sich der V-Wert um ΔV ändert, dann kann a um den Betrag $0.583\sqrt{\Delta V}$ reduziert werden, um das 'overshooting' auszugleichen. Die Achsenabschnitte der beiden Geraden sind dann: $a_0' = -a + 0.583\sqrt{\Delta V}$ und $a_1' = a - 0.583\sqrt{\Delta V}$. Erfolgt im obigen Beispiel (mit $\pi_0 = 0.55$) die Testung jeweils in Gruppen von $N=20$ ($n_1=10$ und $n_2=10$) binären Beobachtungen, dann beträgt $\Delta V = (20/4) \cdot 0.247 = 1.24$. Für $\alpha = \beta = 0.05$ sind somit die Achsenabschnitte -10.12 und $+10.12$ zu nehmen.

9 Simulationsergebnisse

Um die Güte der Dreieckspläne zu überprüfen, wurden Bernoulli- und normalverteilte Zufallsgrößen für unterschiedliche Alternativhypothesen simuliert und damit die entsprechenden Dreieckstests durchgeführt. Die Simulationen wurden jeweils 1000-mal wiederholt und damit die Power und ASN für die Nullhypothese und Alternative ermittelt. Zum Vergleich ist auch der Stichprobenumfang (n_{fix} bei einer Stichprobe bzw. der Gesamtstichprobenumfang N_{fix} bei 2 Stichproben) angegeben, der bei Testung mit fester Stichprobenzahl erforderlich ist.

Die Ergebnisse zeigen, daß auch bei größeren Werten von θ_1 die Signifikanzniveaus von α und β sehr gut eingehalten werden. Der mittlere Stichprobenumfang beim Sequenztest ist oft ca. halb so groß wie er bei einem Test mit festem Stichprobenumfang erforderlich wäre. Dies demonstriert eindrucksvoll den Vorteil des Dreieckstest. Nach Angaben bei Whitehead [7,8]

ist der Test optimal für $\theta = \frac{1}{2}\theta_1$ und besitzt auch für $\theta = 0$ bzw. θ_1 einen ASN-Werte, der nur unwesentlich größer als der SPRT ist (der ja für diese Parameterwerte optimal ist). Die Simulationen haben außerdem die links-steile Verteilung von n (Stichprobenumfang bis zur Entscheidung) bestätigt. In 60-70% der Simulationen war n kleiner oder gleich der ASN.

Simulation von Sequenzttests für eine Stichprobe binärer Werte mit $\text{Pr}(1)=\pi$

$H_0: \pi=\pi_0; H_1: \pi=\pi_1 ; \alpha = 0.05; \beta = 0.2$

π_0	π_1	$P(\pi_0)$	ASN(π_0)	$P(\pi_1)$	ASN(π_1)	n_{fix}
0.1	0.15	0.046	172	0.849	185	253
0.1	0.2	0.068	56	0.878	53	69
0.1	0.3	0.067	20	0.932	16	20
0.2	0.3	0.057	72	0.859	77	109
0.2	0.4	0.066	22	0.889	23	29
0.3	0.4	0.044	82	0.817	95	136
0.3	0.5	0.050	23	0.842	26	35
0.4	0.5	0.057	86	0.827	104	151
0.5	0.6	0.039	80	0.800	101	153
0.5	0.7	0.042	20	0.796	25	37

Simulation von Sequenzttests für den Erwartungswert normal verteilter Werte

$H_0: \theta=\mu/\sigma = 0; H_1: \theta=\mu/\sigma = \theta_1 \alpha = 0.05; \beta = 0.2$

θ_1	$P(0)$	ASN(0)	$P(\theta_1)$	ASN(θ_1)	n_{fix}
0.1	0.045	329	0.828	420	619
0.2	0.041	86	0.825	108	155
0.3	0.051	39	0.816	50	69
0.4	0.050	23	0.823	29	39
0.5	0.055	15	0.793	21	25

Simulation von Sequenzttests für zwei Stichproben mit binären Werten

$H_0: \pi_1=\pi_2; H_1: \pi_2>\pi_1 ; \alpha = 0.05; \beta = 0.2$

π_1	π_2	$P(H_0)$	ASN(H_0)	$P(H_1)$	ASN(H_1)	N_{fix}
0.1	0.2	0.046	243	0.810	207	310
0.1	0.3	0.051	92	0.805	67	93
0.1	0.4	0.048	57	0.816	35	46

0.3	0.4	0.054	330	0.782	374	557
0.3	0.5	0.037	93	0.787	100	143
0.3	0.6	0.049	43	0.829	46	62
0.5	0.6	0.052	334	0.816	402	606
0.5	0.7	0.049	78	0.801	101	143
0.5	0.8	0.053	32	0.810	43	57

Simulation von Sequenztests für zwei Stichproben mit normal verteilten Werten

$$H_0: \mu_1 = \mu_2; H_1: \mu_2 > \mu_1; \sigma = 1; \alpha = 0.05; \beta = 0.2$$

$\mu_2 - \mu_1$	$P(H_0)$	ASN(H_0)	$P(H_1)$	ASN(H_1)	N_{fix}
0.2	0.046	340	0.811	418	619
0.3	0.048	152	0.830	187	275
0.4	0.040	85	0.811	110	155
0.5	0.051	56	0.837	71	99
0.6	0.050	40	0.799	50	69
0.7	0.056	30	0.826	37	51
0.8	0.043	24	0.825	24	39
0.9	0.052	19	0.831	24	31
1.0	0.043	16	0.823	20	25

10 Das 3-Entscheidungs-Problem (zweiseitige und Äquivalenztestung)

Bei Tests mit festem Stichprobenumfängen werden zweiseitige Alternativen mit dem Betrag oder Quadrat einer Teststatistik und dem nominellen Niveau von $\alpha/2$ für jede der beiden Abweichungen getestet. Bei Sequenztests ist eine differenziertere Behandlung des Problems als Entscheidungsproblem zwischen den drei Hypothesen:

$$H_0: \theta = 0; H_{1-}: \theta = \theta_{1-} < 0 \text{ und } H_{1+}: \theta = \theta_{1+} > 0$$

möglich. Die beiden Werte θ_{1-} und θ_{1+} müssen dabei nicht symmetrisch um 0 liegen. Dies ist praktisch von Bedeutung, insbesondere beim Vergleich zweier Wahrscheinlichkeiten, bei dem als Testparameter θ die log-odds-ratio genommen wird. θ -Werte symmetrisch um 0 entsprechen dann oft nicht den gewünschten Abweichungen in den π -Werten. Eine ähnliche Situation ergibt sich bei der Bioäquivalenztestung, bei der eine untere Abweichung von 80% und eine obere von 120% mit dem Logarithmus des AUC-Quotienten der beiden Zubereitungen erkannt werden soll.

Sequentiell wird das 3-Entscheidungs-Problem durch zwei Sequentialpläne gelöst. Der eine Sequentialplan SP_+ entscheidet zwischen der Nullhypothese und der Alternative H_{1+} , der

zweite Plan SP₋ zwischen der Nullhypothese und der Alternative H₁₋. Soll für die irrtümliche Ablehnung der Nullhypothese das Niveau α eingehalten werden, dann wird für beide Sequentialpläne das Niveau $\alpha/2$ vorgegeben. Für die irrtümliche Ablehnung der jeweiligen Alternative wird bei beiden Plänen das Niveau β vorgegeben.

In Abb. 5 ist ein doppelter SPRT zum Test der Nullhypothese: $\pi_1=\pi_2$ gegen die Alternative H₁₋: $\pi_1<\pi_2$ ($\pi_1=0.4$ und $\pi_2=0.5$) und die Alternative H₁₊: $\pi_1>\pi_2$ ($\pi_1=0.6$ und $\pi_2=0.5$) gezeigt, in Abb. 6 der entsprechend Dreieckstest. Für beide Pläne sind jeweils $\alpha=0.025$ und $\beta=0.025$ vorgegeben.

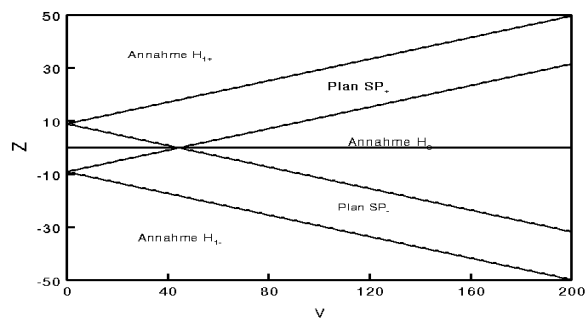


Abb.5 Offener Plan für das Dreieckscheidungsproblem

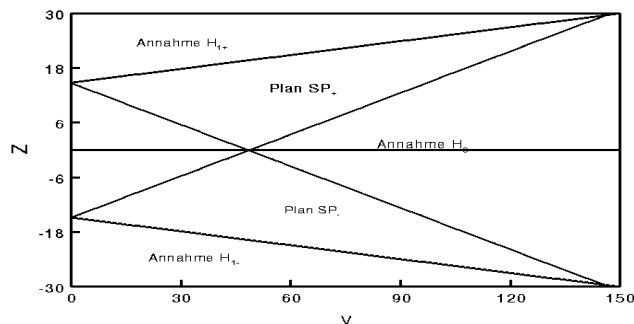


Abb. 6 Geschlossener Plan für das Dreieckscheidungsproblem

Die Datenerhebung wird solange fortgesetzt, wie sich der Pfad in einem der Fortsetzungsbereiche befindet. Erreicht oder überschreitet er die obere Grenze von SP₊, dann wird die Alternative H₁₊ angenommen; erreicht oder überschreitet er die untere Grenze von SP₋, dann wird H₁₋ angenommen. Erreicht oder überschreitet er die untere Grenze von SP₊ oder die obere Grenze von SP₋, dann wird die Nullhypothese angenommen. Die Wahrscheinlichkeit, bei Gültigkeit von H₀ eine der Alternativen anzunehmen, ist höchstens $\alpha=0.05$; die Wahrscheinlichkeit, die Nullhypothese anzunehmen, wenn eine der Alternativen gilt, ist höchstens $\beta=0.025$.

Ein Problem kann entstehen, wenn sich der Pfad noch im Überdeckungsbereich beider

Fortsetzungsregionen befindet und vom Fortsetzungsbereich des einen Plans in den des anderen wechselt; z. B. vom Fortsetzungsbereich von SP_+ in den von SP_- . Sobel und Wald haben vorgeschlagen, dann nur noch den Sequenztest fortzuführen, in dessen Fortsetzungsbereich der Pfad gewechselt ist (im Beispiel also den Plan SP_-). Es wird dann die Entscheidung getroffen, zu der dieser Plan führt. Erreicht im Beispiel der Pfad, nachdem er in den Fortsetzungsbereich von SP_- gewechselt ist, die obere Grenze dieses Bereichs, dann wird für H_0 entschieden, auch wenn der Pfad wieder zurück in den Fortsetzungsbereich von SP_+ wechselt. Eine Entscheidung für H_{1+} ist dann nicht mehr möglich.

Erreicht der Pfad den Annahmebereich für H_0 , dann kann mit Wahrscheinlichkeit $1-\beta$ behauptet werden, daß der Parameter θ zwischen θ_{1-} und θ_{1+} liegt; d.h. die Wahrscheinlichkeit, den Annahmebereich für H_0 zu erreichen, wenn θ einen dieser alternativen Werte oder einen noch extremeren Wert besitzt, ist jeweils höchstens β . Die Annahme von H_0 ist somit gleichbedeutend mit der Annahme der Äquivalenz, wenn θ den Unterschied zwischen zwei Parametern (Wahrscheinlichkeiten, Mittelwerte) ausdrückt und als Äquivalenzbereich das Intervall $(\theta_{1-}, \theta_{1+})$ vorgegeben ist. Ein sequentielles 3-Entscheidungs-Verfahren liefert so einen Äquivalenztest.

11 α - und β -spending Funktionen

Ein Sequenzplan ist durch die Vorgabe des Fortsetzungsbereichs festgelegt, der von der Annahmegrenze für H_1 und von der für H_0 begrenzt ist. Erreicht ein Pfad bei einem Stichprobenumfang n (bzw. einem Wert v der Statistik V) eine dieser Grenzen, dann wird die Stichprobenerhebung abgeschlossen und die entsprechende Entscheidung getroffen. Der Umfang n (bzw. v), bei dem dieses geschieht, und die Grenze, die erreicht wird, sind Realisationen zufälliger Ereignisse N (bzw. V) und E , deren Verteilung von θ abhängt. Die Zufallsgröße E kann die beiden Werte 0 (Annahme von H_0) und 1 (Annahme von H_1) annehmen. Die Wahrscheinlichkeit, daß ein Pfad die Annahmegrenze für H_1 bei einem $N \leq n$ erreicht, wenn H_0 gilt, gibt das nominelle Signifikanzniveau $\alpha(n)$ an, das von der Zufallsgröße Z bis zum Stichprobenumfang n (bzw. v) erreicht werden muß, wenn H_1 bis dahin angenommen werden soll. Dieses Niveau nimmt mit n monoton von 0 auf das vorgegebene Niveau α zu. Es ist bei geschlossenen Plänen üblich, statt des Stichprobenumfangs n (bzw. v) den Anteil am maximalen Umfang N_{\max} (bzw. V_{\max}) anzugeben; d.h. $t=n/N_{\max}$ (bzw. $t=v/V_{\max}$). Die Funktion

$$\alpha(t) = \Pr(\text{Annahme } H_1 \text{ für } T \leq t | H_0)$$

nennt man die **α -spending Funktion**. Dabei bedeutet T die Zufallsgröße N/N_{\max} (bzw. V/V_{\max}). Die α -spending Funktion wurde von Lan und DeMets [5] eingeführt. Sie ist für $0 \leq t \leq 1$ definiert und nimmt in diesem Intervall monoton von 0 auf α zu. Diese Funktion kann als Entscheidungsgrenze zur Annahme von H_1 benutzt werden. Es sind dann statt der Werte z_m , die sich für die Teststatistik Z_m mit den Stichprobenwerten x_1, \dots, x_m ergeben, die Werte der entsprechenden Signifikanzwahrscheinlichkeiten $s_m = \Pr(Z_m > z_m | H_0)$ über m bzw. $t_m = m/N_{\max}$ als Pfade in ein Koordinatensystem einzutragen. Erreicht oder überschreitet der Pfad (t_m, s_m) für ein $t=t_n$ den entsprechenden Wert $\alpha(t_n)$ der α -spending Funktion, dann ist die Stichprobenerhebung zu beenden und die Alternativhypothese H_1 anzunehmen. In Abb. 7

sind für einen standardisierten Dreiecksplan mit $\alpha=\beta=0.05$ die $\alpha(t)$ -Werte für ausgewählte Stellen v angegeben, in Abb. 8 die zugehörige α -spending Funktion.

Analog läßt sich auch die β -spending Funktion definieren und als Entscheidungsgrenze für die Annahme von H_0 benutzen: $\beta(t) = \Pr(\text{Annahme } H_0 \text{ für } T \leq t | H_1)$. Diese Funktion nimmt im Intervall $0 \leq t \leq 1$ monoton von 0 auf den vorgegebenen Wert β zu.

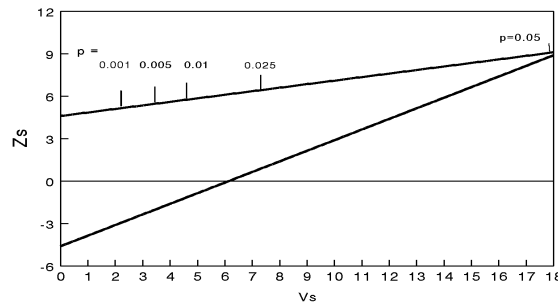


Abb. 7. Das nominelle Signifikanzniveau $\alpha(t)$ des Dreiecksplans

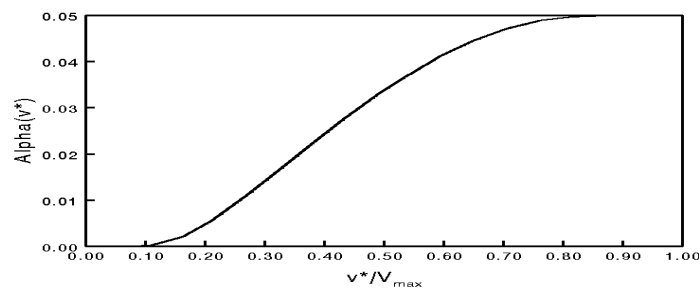


Abb. 8 Die α -spending Funktion $\alpha(t)$ des Dreiecksplans

In Abb. 9 sind simulierte β -spending Funktionen für den 2-Stichprobenvergleich binärer Größen mit dem Dreieckstest gezeigt. Die Ergebnisse beruhen auf Simulationen von jeweils 1000 Pfaden.

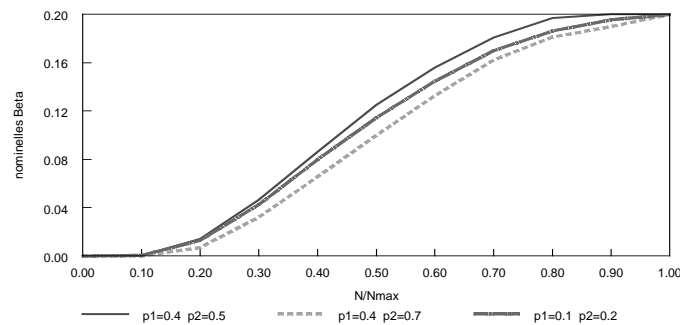


Abb. 9 Simulierte β -spending Funktionen für den 2-Stichprobenvergleich binärer Größen

12 Konfidenzintervalle für θ

Nach Abschluß einer Studie möchte man nicht nur eine Testentscheidung für eine der beiden Hypothesen treffen, sondern auch über den Parameter θ etwas aussagen; z.B. ein Konfidenzintervall für θ angeben. Bei fest vorgegebenem Stichprobenumfang ist das folgende Vorgehen angebracht. Soll z.B. bei normal verteilten Stichprobengrößen für den Parameter $\theta = \mu/\sigma$ die Nullhypothese $\theta = 0$ gegen eine Alternative $\theta = \theta_1 > 0$ mit einem Stichprobenumfang n getestet werden, dann wird man die Testgröße $t = \frac{\bar{x}}{s} \sqrt{n}$ nehmen, die t-verteilt mit $n-1$

Freiheitsgraden und dem Nichtzentralitätsparameter $\theta \sqrt{n}$ ist. Wurde mit der Stichprobe der Wert t^* erhalten, dann ist $C(\theta; t^*) = \Pr(t > t^* | \theta) = 1 - F_{t, n-1}(t^* - \theta \sqrt{n})$ eine monotone Funktion von θ (wobei $F_{t, n-1}(\cdot)$ die Verteilungsfunktion der zentralen t-Verteilung mit $n-1$ Freiheitsgraden ist). Die untere Grenze θ_u und obere Grenze θ_o eines $(1-\gamma)$ -Konfidenzintervalls für θ sind Lösungen der beiden Gleichungen: $C(\theta_u; t^*) = \gamma/2$ und $C(\theta_o; t^*) = 1 - \gamma/2$. Diese Lösungen lauten explizit:

$$\theta_u = \frac{t^* - t_{1-\gamma/2; n-1}}{\sqrt{n}} = \frac{\bar{x}}{s} - \frac{t_{1-\gamma/2; n-1}}{\sqrt{n}} \quad \text{und} \quad \theta_o = \frac{t^* + t_{1-\gamma/2; n-1}}{\sqrt{n}} = \frac{\bar{x}}{s} + \frac{t_{1-\gamma/2; n-1}}{\sqrt{n}}$$

wobei $t_{1-\gamma/2; n-1}$ die $(1-\gamma/2)$ -Quantile der zentralen t-Verteilung mit $n-1$ Freiheitsgraden ist. Die Funktion $C(\theta; t^*)$ kann als 'Konfidenzfunktion' oder 'Konfidenzverteilung' bezeichnet werden. In der Bayesianischen Statistik, bei der θ als eine Zufallsgröße aufgefaßt wird, ist $C(\theta; t^*)$ die a posteriori Verteilung von θ bei nichtinformativer a priori Verteilung. Diese kann auch als 'Fiduzialverteilung' für den Parameter θ (im Sinne von R.A. Fisher) interpretiert werden. Der Konfidenzbereich entspricht dem 'credible interval'.

Bei Sequentialtests stehen bei Abschluß zwei Informationen zur Verfügung: die Entscheidung e , die entweder in der Annahme von H_0 ($e=0$) oder in der Annahme von H_1 ($e=1$) bestehen kann, und der Stichprobenumfang n (bzw. der Wert v der Statistik V), bei dem es zu dieser Entscheidung kam. Der Wert e ist Realisation einer Zufallsgröße E und n Realisation einer Zufallsgröße N . Ihre Verteilung hängt von θ ab. Für gegebene Werte e und n ist die Konfidenzfunktion $C(\theta; n, e)$ definiert als:

$$C(\theta; n, 1) = \Pr(N \leq n \text{ und } e=1; \theta) \quad \text{und} \quad C(\theta; n, 0) = 1 - \Pr(N \leq n \text{ und } e=0; \theta)$$

$C(\theta; n, e)$ steigt für gegebenes e und n von 0 (für $\theta \rightarrow -\infty$) auf 1 (für $\theta \rightarrow \infty$) monoton an. Damit lassen sich die Grenzen eines $(1-\gamma)$ -Konfidenzintervalls für θ berechnen. Sie sind Lösung der Gleichungen:

$$C(\theta_u; n, e) = \gamma/2 \quad \text{und} \quad C(\theta_o; n, e) = 1 - \gamma/2$$

Als Schätzwert für θ kann der Wert θ_M genommen werden, für den gilt: $C(\theta_M; n, e) = 1/2$.

In Abb. 10 sind die Funktionen $C(\theta; n, e)$ für den Dreieckstest zum Vergleich der Wahrscheinlichkeiten π_1 und π_2 zweier Stichproben binärer Daten für verschiedene Werte

von $t=n/N_{\max}$ sowie für $e=1$ (Abb.10a) und $e=0$ (Abb.10b) gezeigt. Als Nullhypothese wurde $\pi_1=\pi_2=0.4$, als Alternativhypothese $\pi_1=0.4$ und $\pi_2=0.5$ angenommen. Vorgegeben sind die Niveaus $\alpha=0.05$ und $\beta=0.2$. Die Kurven wurden aus der Simulation von jeweils 1000 Pfaden ermittelt. Auf der Abszisse ist θ/θ_1 aufgetragen. Die drei Kurven entsprechen den Stichprobengrößen $t=n/N_{\max}=0.3, 0.5$ und 0.7 .

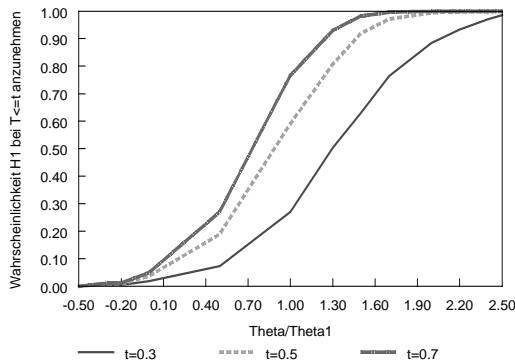


Abb. 10a: $C(\theta; t, 1)$

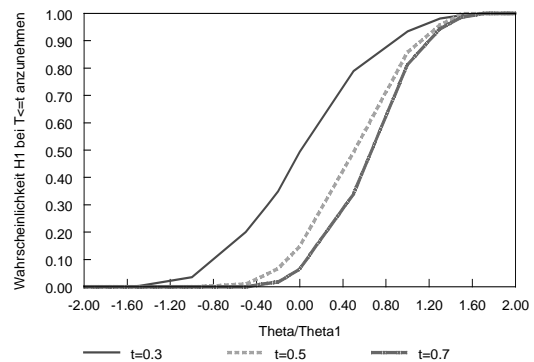


Abb. 10b: $C(\theta; t, 0)$

Abb. 10: Simulierte Konfidenzfunktionen für Dreieckstests der Hypothese $\pi_1=\pi_2$ gegen $\pi_1<\pi_2$
 $\pi_1=0.4, \pi_2=0.5, \alpha=0.05, \beta=0.2$

Die mit diesen Konfidenzfunktionen ermittelten 95%-Konfidenzintervalle und die entsprechenden Schätzwerte θ_M sind in der folgenden Tabelle gezeigt. Man beachte, daß für diese Fragestellung $\theta=\ln(\pi(1-\pi_1)/(\pi_1(1-\pi)))$ ist, wobei π die Wahrscheinlichkeit für das Ereignis ($x=1$) für die zweite Stichprobe und π_1 für die erste Stichprobe bezeichnet. In der Tabelle sind auch die entsprechenden Konfidenzintervalle und der Schätzwert für π_2 angegeben, die unter der Annahme $\pi_1=0.4$ gelten.

Annahme von H_1 ($e=1$)						
t=0.3	$\theta_u=0.162$	$\theta_o=0.973$	$\theta_M=0.527$	$\pi_u=0.439$	$\pi_o=0.638$	$\pi_M=0.530$
t=0.5	$\theta_u=0.000$	$\theta_o=0.689$	$\theta_M=0.405$	$\pi_u=0.400$	$\pi_o=0.570$	$\pi_M=0.500$
t=0.7	$\theta_u=-0.061$	$\theta_o=0.588$	$\theta_M=0.304$	$\pi_u=0.385$	$\pi_o=0.545$	$\pi_M=0.474$

Annahme von H_0 ($e=0$)						
t=0.3	$\theta_u=-0.487$	$\theta_o=0.487$	$\theta_M=0.000$	$\pi_u=0.290$	$\pi_o=0.520$	$\pi_M=0.400$
t=0.5	$\theta_u=-0.162$	$\theta_o=0.689$	$\theta_M=0.203$	$\pi_u=0.362$	$\pi_o=0.570$	$\pi_M=0.449$
t=0.7	$\theta_u=-0.041$	$\theta_o=0.588$	$\theta_M=0.243$	$\pi_u=0.390$	$\pi_o=0.545$	$\pi_M=0.460$

Man erkennt, daß bei Annahme von H_1 die Schätzwerte θ_M und die Grenzen des Konfidenzintervalls um so größer sind, je früher die Entscheidung für H_1 fällt (je kleiner t

ist). Bei Annahme von H_0 ist es umgekehrt. Bei einer früheren Entscheidung für H_0 sind die Schätzwerte und Konfidenzgrenzen für θ kleiner als bei einer späteren Entscheidung. Eine Reihenentwicklung für die Konfidenzfunktion des Dreieckstests ist in Abschnitt 8 gegeben. Die Funktion wurde dort als $Q(e,v;\theta)$ bezeichnet.

13 Programme für Sequenztests

Whitehead hat ein Programmpaket (PEST II) entwickelt, mit dem auf der Basis von Z und V Pläne für SPRT und Dreieckspläne erstellt, deren Power- und ASN-Funktion sowie Konfidenzintervalle für θ berechnet werden können. Das Programm, dessen Handhabung allerdings nicht ganz einfach ist, ist zu beziehen vom: Department of Applied Statistics, University of Reading, Whiteknights, PO Box 217, Reading RG6 2AN, UK (nach Angabe in [9]). Ein wesentlich einfacheres Programmpaket, mit dem Dreieckspläne für 2-Stichprobenvergleiche erstellt und mit diesen Plänen ein sequentielles Datenmonitoring durchgeführt werden kann, wurde von der Fa. BIORAT erstellt (Programm TRIQ) und ist zu beziehen über: BIORAT GmbH, Zentrum für Statistische Beratung und Datenverarbeitung, Joachim-Jungius-Str. 9, 18059 Rostock. Das Programm TRIQ enthält keine Module für die Analyse nach Abschluß der Datenerhebung, kann dafür aber sehr einfach auch von weniger geübten Benutzern implementiert und bedient werden. Für gruppensequentielle Tests auf der Basis der α -spending und β -spending Funktionen wurde von C.R. Mehta und N.R. Patel das Programmsystem EaSt entwickelt. Dieses ist über die Cytel Software Corporation, 675 Massachusetts Avenue Cambridge, MA zu beziehen.

Literatur:

1. Armitage, P.: Some sequential tests of Student's hypothesis. Journ. Roy. Statist. Soc., Supp.; 9, 250-263, 1947
2. Armitage, P.: Sequential medical trials. Blackwell Scientific Publications, Oxford 1960
3. Bauer, P., Scheiber, V. und Wohlzogen, F. X.: Sequentielle statistische Verfahren (Reihe biometrie (Hrsg.: R. J. Lorenz und J. Vollmar). Gustav Fischer Verlag, Stuttgart, New York 1986
4. Cox, D. R.: Sequential tests for composite hypotheses. Proc. Camb. Phil. Soc. 48, 290-299, 1952
5. Lan, K.K.G. and DeMets, D.L.: Discrete sequential boundaries für clinical trials. Biometrika 70, 659-663, 1983
6. Wald, A.: Sequential analysis. John Wiley & Sons, Inc., New York, London, Sidney 1947
7. Wetherill, G. B.: Sequential methods in statistics. Monographs on Applied Probability and Statistics. Chapman and Hall, London 1975
8. Whitehead, J.: The design and analysis of sequential clinical trials, first edition. Ellis Horwood Ltd., Chichester 1983
9. Whitehead, J.: The design and analysis of sequential clinical trials, second edition. Ellis Horwood Ltd., Chichester 1992