

Expertise

**Berechnung des Stichprobenumfangs für die
attributive Qualitätskontrolle bei seltenen Ereignissen
unter Nutzung von Vorinformationen**

- ein Bayesscher Zugang -

Statistik-Service • Dr. Johannes Gladitz

Für die (fiktive) ABC GmbH

Januar 2007



Inhaltsverzeichnis

Einleitung	3
Das statistische Problem	4
Fiduzialintervalle	6
Eine Bayessche Formulierung des Entscheidungsproblems	10
Fallzahlbestimmung unter Einbeziehung von Vorinformationen	13
Zusammenfassung	19
Literatur	20

Bemerkung: Aufgrund dieser Expertise konnte die Zulassung eines neuen Medizinproduktes durch die zuständige Behörde – nach anfänglicher Skepsis – erreicht werden. Die Qualitätsprüfung ist nur durch Zerstörung der hergestellten Produkte möglich. Anstatt einer Stichprobe von 2.995 zu zerstörenden Transplantaten konnte die nötige Fallzahl durch die Einbeziehung von Vorinformationen unter einem Bayesschen Zugang auf eine Fallzahl von 426 Proben reduziert werden, bei garantierter Sicherheit.

Einleitung

Für die Zulassung eines neuen Transplantates ist es erforderlich, den Nachweis zu erbringen, dass bei der Produktion ein vorgegebener Sterilitäts-Sicherheits-Schwellwert (SAL=Sterility Assurance Level) nicht überschritten wird.

Diese Expertise beschreibt eine Methode zur Bestimmung des Stichprobenumfanges zum Nachweis eines vorgegebenen SAL von 0,001 bei Verwendung eines Bayesschen Zuganges unter Nutzung von Vorinformationen, wenn im Prüfungsprozess kein einziger lebensfähiger Keim nachgewiesen wird.

Ziel der technischen Prüfung ist der Nachweis, dass die Firma ABC GmbH in der Lage ist, das neue Transplantat unter höchsten Sterilitätsanforderungen zu produzieren. Mit einer Sicherheit von 95% darf unter 1000 produzierten Produkten im Mittel höchstens ein kontaminiertes Produkt auftreten. Die Produktionsbedingungen müssen also gewährleisten, dass die Auftretenswahrscheinlichkeit einer Kontamination mit einer Sicherheit von 95% nicht höher als 0,1% ist.

Die Prüfung auf eine vorliegende Kontamination ist nur durch eine Zerstörung des Transplantats möglich. Das Ergebnis der Prüfung Y kann nur zwei Ausgänge haben: 0 = keine Kontamination und 1 = Kontamination mit einem lebensfähigen Keim. Die Wahrscheinlichkeit allgemein bezeichnen wir (groß) P , die unbekannte abzuschätzende Kontaminationswahrscheinlichkeit bezeichnen wir mit (klein) p . Es gilt also

$$P(Y=1) = p \quad \text{mit} \quad 0 \leq p \leq 1$$

und p ist unbekannt. Der Parameter p ist die unserem Produktionsprozess inhärente Ausfallwahrscheinlichkeit. Ein Produktausfall ist hier das Auftreten eines lebensfähigen Keims. Wir müssen statistisch abgesichert nachweisen, dass $p \leq 0,001$.

Das vorliegende Problem ist eine attributive Qualitätskontrolle mit zerstörender Produktprüfung bei sehr seltenen Ausfällen. Die Produktion der in Frage stehenden Transplantate ist sehr teuer und der SAL-Nachweis nach den Prinzipien der klassischen statistischen Testtheorie würde die Kosten vor Produktionsaufnahme exorbitant erhöhen.

Die Firma ABC GmbH produziert schon seit einigen Jahren vergleichbare Produkte, hat Erfahrungen mit hochsterilen Produktionsumgebungen und verfügt aus den letzten Jahren über eine Stichprobe von insgesamt 1.046 zerstörend geprüften Transplantaten, bei denen nicht ein einziges Mal ein lebensfähiger Keim nachgewiesen wurde. Gleichzeitig wird ein hochsteriles Ausgangsprodukt mit einem SAL von 10^{-6} verwendet und die ABC GmbH verfügt über das nötige Knowhow zur Produktion von Transplantaten in keimfreien Umgebungen.

Dieses Wissen werden wir quantifiziert in unseren Entscheidungsprozess einbeziehen unter Nutzung eines Bayesschen Entscheidungsansatzes (Gladitz 1976, 1994).

Das statistische Problem

Nehmen wir an, wir haben n Transplantate produziert und alle werden auf lebensfähige Keime überprüft. Dann genügt die Wahrscheinlichkeit dafür, dass die (zufällige) Anzahl K der kontaminierten Transplantate genau k beträgt, einer Binomialverteilung $B(n,p)$ mit:

$$P(K = k) = \binom{n}{k} p^k (1-p)^{n-k} \quad (1)$$

Wenn wir $K = 0$, d.h. keine einzige Kontamination beobachtet haben, wollen wir eine Abschätzung für das unbekannte p nach oben angeben, die mit einer Sicherheit von 95% richtig ist. Eine solche Abschätzung liefert hier ein einseitiges 95% Konfidenzintervall mit den Grenzen $(0 ; p_0)$.

Aus den Formeln von Clopper und Pearson (1934) ergibt sich für $p_0 = 0,001$ ein minimaler Stichprobenumfang $n = 2.995$.

Nach dem klassischen frequentistischen Konzept der Statistik müsste man also knapp 3.000 Produkte herstellen und zerstörend prüfen, um dann bei keiner einzigen Kontamination auf einen $SAL < 0,001$ mit 95% Sicherheit schließen zu können.

Man kann das Konfidenzintervall hier auch einfach berechnen. In unserem Fall vereinfacht sich die Formel (1), wenn wir $k=0$ setzen:

$$P(K = 0) = \binom{n}{0} p^0 (1-p)^{n-0} = (1-p)^n \quad (2)$$

Betrachten wir die Formel (2) als Funktion von p für gegebenes n , erhalten wir die Likelihoodfunktion. Ein exaktes 95% Konfidenzintervall $(0 ; p_0)$ für den unbekannten Parameter p erhalten wir dann aus der Lösung der Gleichung:

$$P(K = 0) = L(p) = (1-p)^n = 0,05 \quad (3)$$

Nach p aufgelöst ergibt das für unsere gesuchte Obergrenze p_0 :

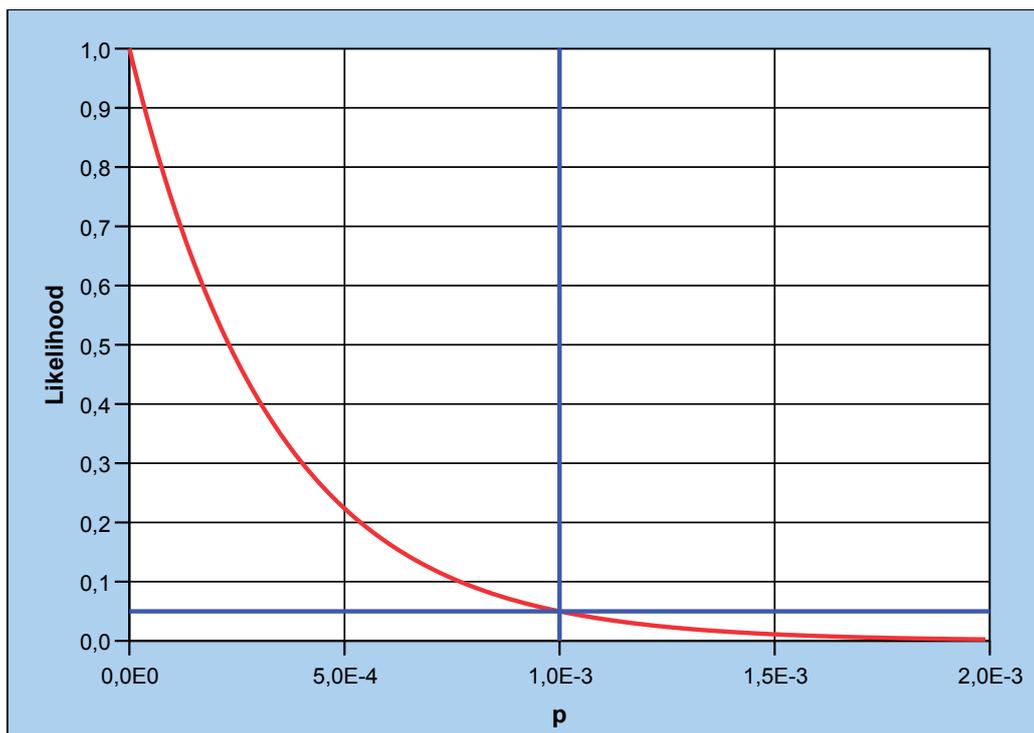
$$p_0 = 1 - 0,05^{(1/n)} \quad (4)$$

Damit ist die Wahrscheinlichkeit, einen Wert von $K = 0$ zu erhalten, wenn in Wirklichkeit $p > p_0$ ist, kleiner als 0,05. Das wäre die Irrtumswahrscheinlichkeit. Wenn wir uns also nach einer Beobachtung von $K = 0$ für $p \leq p_0$ entscheiden, treffen wir in 95% aller möglichen Fälle eine richtige Entscheidung, denn in weniger als 5% aller Fälle kann ein Ergebnis $K=0$ herauskommen, wenn in Wirklichkeit $p > p_0$ wäre.

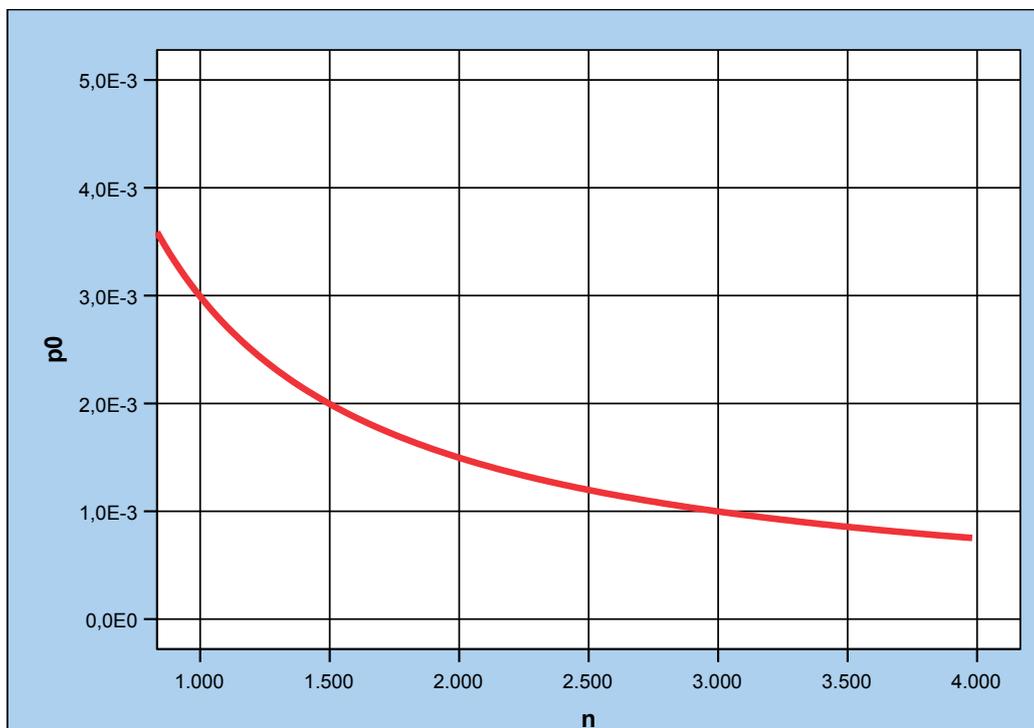
Damit ist das Intervall $(0 ; p_0)$ ein (frequentistisches) 95% Konfidenzintervall.

Auf der folgenden Seite stellen wir die Zusammenhänge grafisch dar.

Verlauf der Likelihoodfunktion für $n=2.995$



Die kritische Obergrenze p_0 in Abhängigkeit von der Fallzahl n



Etwa bei $n = 3.000$ erreicht die kritischen Obergrenze $SAL = p_0$ den gewünschten Wert von 10^{-3} .

Der hier beschriebene Standardansatz berücksichtigt jedoch nicht alle vorliegenden Informationen.

Fiduzialintervalle

Die Likelihoodfunktion (3) bewertet die Passfähigkeit der möglichen Parameterwerte p zum Ergebnis unserer Stichprobe $K=0$. Heute ist das Maximum-Likelihood-Konzept allgemein akzeptiert. Man wählt einen unbekannt Parameter \hat{p} so, dass dem beobachteten Wert $K=k$ in der Stichprobe die maximale Plausibilität zukommt, d.h. die Likelihoodfunktion (Stichprobenfunktion) für \hat{p} maximiert wird. Bei uns ergibt sich $\hat{p} = 0$.

Die Likelihoodfunktion enthält also die Plausibilität der verschiedenen Parameterwerte in Bezug zum Ausgang unseres Experiments. Sie hat die Form einer Dichtefunktion über dem Parameterraum mit einem Schönheitsfehler: Sie ist nicht notwendig normiert, d.h. die Fläche unter der Likelihoodfunktion muss nicht notwendig 1 sein.

Wenn das Integral über der Likelihoodfunktion existiert, kann man diese damit normieren. Dabei ergibt sich formal eine Dichtefunktion über dem Parameterraum. Die zugehörige Verteilung nennt man Fiduzialverteilung. Sie wurde 1935 von Ronald A. Fisher eingeführt. Man könnte die Fiduzialverteilung als strukturelle Parameterverteilung bezeichnen, die sich unmittelbar aus der Stichprobenverteilung als normierte Likelihoodfunktion ergibt. Damit ergibt sich in unserem Fall eine Verteilungsfunktion über dem Intervall $(0;1)$, welche es erlaubt, die Plausibilität der möglichen Parameterwerte bzw. von Parameterintervallen mit Hilfe von Wahrscheinlichkeiten auszudrücken.

Allgemein hat die Likelihoodfunktion für eine binomiale Stichprobenverteilung die Form

$$P(K = k|p, n) = L(p|k, n) = \binom{n}{k} p^k (1-p)^{n-k} \quad (5)$$

Daraus ergibt sich die zugehörige Fiduzialverteilung zu

$$\begin{aligned} F(p|k, n) &= \frac{p^k (1-p)^{n-k}}{\int_0^1 p^k (1-p)^{n-k} dp} = \frac{\Gamma(n+2)}{\Gamma(k+1)\Gamma(n-k+1)} p^k (1-p)^{n-k} \\ &= (n+1) \binom{n}{k} p^k (1-p)^{n-k} \end{aligned} \quad (6)$$

Γ ist die Gammafunktion. Für diese gilt bekanntermaßen $\Gamma(x+1) = x\Gamma(x)$ (mit $\Gamma(1) = 1$), woraus für ganzzahlige n die Beziehung $\Gamma(n) = (n-1)!$ folgt.

Die Funktion (6) ist die Dichte einer Betaverteilung mit den Strukturparametern $k+1$ und $n-k+1$. Wir verwenden die Bezeichnung $\text{Beta}(\alpha, \beta)$ hier mit $\alpha = k+1$ und $\beta = n-k+1$.

Die Fiduzialverteilung enthält unser unscharfes Wissen über den unbekannt Parameter p nach dem Experiment. Angewendet auf unser Problem erhalten wir ($k=0$)

$$F(p|k=0, n) = \frac{\Gamma(n+2)}{\Gamma(1)\Gamma(n+1)} (1-p)^n = (n+1)(1-p)^n \quad (7)$$

was hier die Dichtefunktion einer Beta-Verteilung mit $\alpha = 1$ und $\beta = n+1$ liefert. Dabei handelt es sich um eine streng monoton abfallende Funktion, die für $p=0$ den Wert $n+1$ und für $p=1$ den Wert 0 annimmt. Ein Konfidenzintervall wählt man jetzt so, dass die Fläche unter der Fiduzialdichte genau 95% ergibt und gleichzeitig das Intervall so kurz wie möglich wird. Daraus folgt in unserem Fall, dass wir beginnend mit der Untergrenze 0 die Obergrenze genau so weit verschieben, bis die Fläche unter der Fiduzialdichte genau 0,95 wird. Diese Fläche entspricht aber der kumulativen Summenfunktion der Beta-Verteilung mit den Parametern $\alpha = 1$ und $\beta = n+1$.

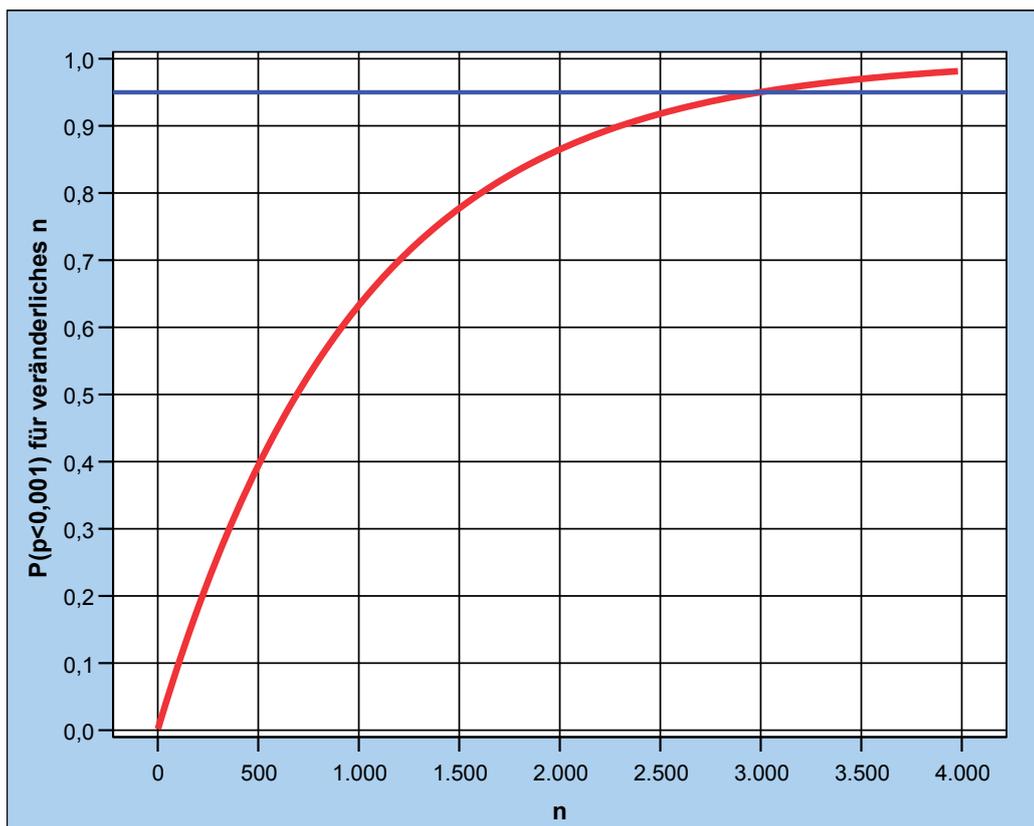
Bei gegebener Fallzahl n und bei $K=0$ ergibt sich somit die Obergrenze p_0 aus der inversen Verteilungsfunktion der Beta-Verteilung mit den Parametern $\alpha = 1$ und $\beta = n+1$:

$$\text{BETAINV}(0,95;1;n+1) = p_0.$$

Diese Funktion findet man u. a. im Office-Programm Excel.

Falls der kritische Wert $SAL=p_0=0,001$ vorgegeben wird, kann man die minimal erforderliche Fallzahl für den Prüfplan berechnen.

Beta-Verteilungsfunktion bei $p=0,001$ mit $\alpha = 1$ und $\beta = n+1$



Das Diagramm zeigt die Veränderung des erreichten Konfidenzniveaus für das Intervall $(0 ; 0,001)$ im Parameterraum in Abhängigkeit von der Fallzahl n bei Verwendung des Fisherschen Fiduzialkonzepts, wenn $K=0$ ist.

Auch hier benötigen wir einen Stichprobenplan etwa vom Umfang 3.000 (für $n=2.994$ wird $p_0 = 0,001$), um die vorgeschriebene Sicherheit zu erlangen.

In unserem Fall kann man p_0 mit einfachen Mitteln direkt bestimmen, und zwar aus der Lösung des folgenden Problems:

$$\int_0^{p_0} (n+1)(1-p)^n dp = 0,95 \quad , \text{ was äquivalent ist mit}$$

$$(n+1) \int_{p_0}^1 (1-p)^n dp = 0,05 \quad . \quad (8)$$

Eine Stammfunktion für das Integral (8) ergibt sich zu

$$(n+1) \frac{-1}{n+1} (1-p)^{n+1} \Big|_{p_0}^1 = (1-p_0)^{n+1} = 0,05 \quad . \quad (9)$$

Als Konfidenzgrenze unter Verwendung des Fiduzialansatzes ergibt sich also fast dieselbe Lösung wie beim frequentistischen Vorgehen (4), nämlich

$$p_0 = 1 - 0,05^{1/(n+1)} \quad . \quad (10)$$

Damit fällt p_0 minimal geringer aus als beim frequentistischen Zugang.

Für $n = 2.995$ ergeben sich folgende Werte für p_0 :

$$\begin{aligned} p_0^{\text{frequentistisch}} &= 0,000999744 \\ p_0^{\text{fiduzial}} &= 0,000999410 \end{aligned}$$

Für praktische Zwecke können die Ergebnisse beider Konzepte als identisch angesehen werden.

Die Wahl des Konzeptes ist eher eine philosophische Frage. In beiden Fällen ist der Ausgangspunkt für die Bestimmung der Fallzahl hier die Likelihoodfunktion. Diese ergibt sich aus der logischen Struktur der Versuchsanordnung und unterliegt keiner Willkür. Bei frequentistischem Ansatz wird indirekt auch der Parameterraum bewertet, indem man die Frage beantwortet, wie groß ist die Wahrscheinlichkeit, ein Ergebnis von $K=0$ zu erhalten, wenn in Wirklichkeit $p=p_0$ wäre. Für variierende Werte p_0 ergeben sich so unterschiedliche Wahrscheinlichkeiten für $K=0$. Das Konfidenzintervall schließt dann diejenigen Parameter aus dem Intervall $[0,1]$ aus, für welche die Auftretenswahrscheinlichkeit von $K=0$ kleiner als 5% wäre. Problematisch bei diesem Zugang ist, dass über K eigentlich keine Ungewissheit mehr besteht, denn $K=0$ ist schon eingetreten. Die Ungewissheit besteht über p .

An dieser Schwäche setzt das Fiduzialkonzept an. Es bewertet den Parameterraum im Hinblick auf seine Passfähigkeit zum gegebenen Versuchsausgang $K=0$ durch eine Wahrscheinlichkeitsverteilung, die ihre Legitimation aus der Struktur der Likelihoodfunktion bezieht. Diese strukturelle Parameterverteilung ist nichts anderes als eine unscharfe Menge (fuzzy set im Sinne von Lotfi A. Zadeh (1965)), deren Zugehörigkeitsfunktion (membership function) die Dichte unserer Betaverteilung darstellt. Likelihoodkonzept und Fuzzy Logic sind heute in der Technik allgemein akzeptierte Entscheidungsansätze. Das Fiduzialkonzept ist eine Kombination aus beiden, und ver-

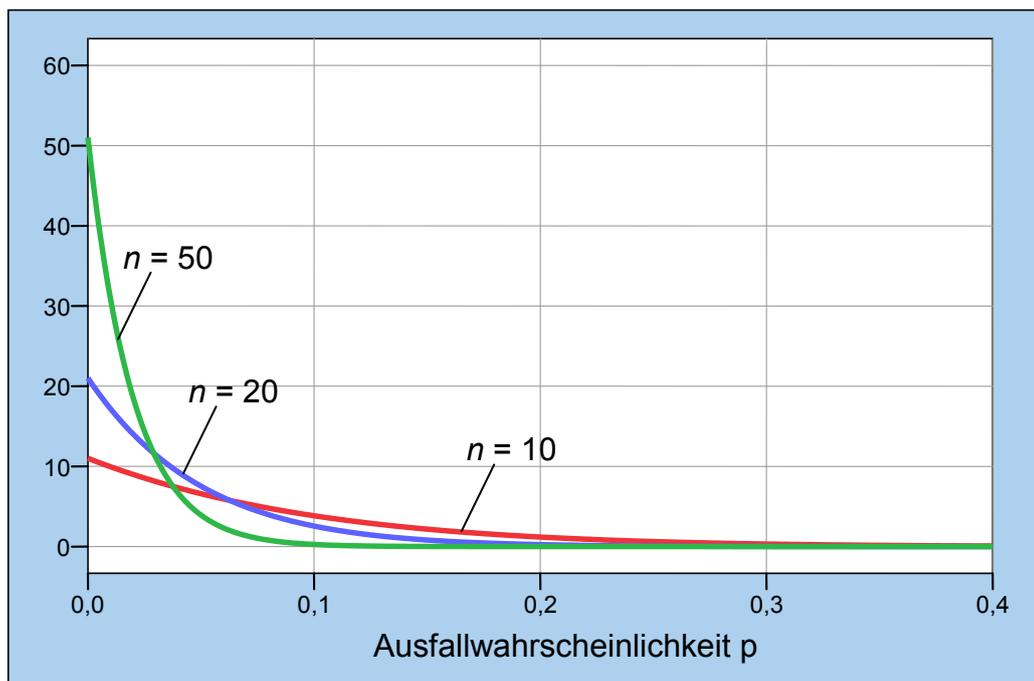
wendet keine rein subjektive Verteilung zur Beschreibung unseres Wissens über das unbekannte p , sondern eine Verteilung, die sich objektiv als strukturelle Parameterverteilung aus der Logik der Versuchsanordnung ergibt.

Aus der Sicht des Gutachters ist das Fiduzialkonzept zur Bewertung unser Wissen über den unbekannt Parameter nach dem Versuch einleuchtender und verständlicher als der frequentistische Zugang zum Konfidenzintervall, dessen heuristische Grundlage die angenommene häufige Wiederholung desselben Versuches ist, die aber praktisch nicht stattfindet.

Auch bei adaptiven oder sequentiellen Entscheidungsprozessen (etwa bei selbstlernenden Systemen) sind Fiduzialverteilungen geeignet, das Wissen über den unbekannt Parameter zu quantifizieren. Gegenüber den oft willkürlich gewählten Zugehörigkeitsfunktionen unscharfer Mengen ergeben sich die Fiduzialverteilungen ganz natürlich.

Das folgende Diagramm zeigt die Veränderung unseres Wissens über die unbekannte Ausfallrate bei verschiedenen Fallzahlen, wenn in jeder Versuchsreihe kein einziger lebensfähiger Keim gefunden wurde.

Fiduzialverteilung für p bei $K=0$ für verschiedene Fallzahlen



Im folgenden Abschnitt werden wir einen Bayesschen Zugang zu unserem Problem vorstellen.

Eine Bayessche Formulierung des Entscheidungsproblems

Bei unserer fiduzialen Betrachtung im vorigen Abschnitt haben wir unser Wissen über den unbekanntem Ausfallparameter p nach der attributiven Qualitätsprüfung durch eine Wahrscheinlichkeitsverteilung über dem Parameterraum beschrieben. Diese heißt Fiduzialverteilung und hatte bei uns die Form einer speziellen Betaverteilung. Sie ergab sich als eine Schlussfolgerung aus dem Likelihoodkonzept.

So wie wir unsere Informationen über p nach dem Experiment in Form einer Verteilung ausgedrückt haben, können wir versuchen, auch unser unscharfes Wissen vor dem eigentlichen Test mit einer geeigneten Wahrscheinlichkeitsdichte zu modellieren und so in den Entscheidungsprozess mit einfließen zu lassen.

Oft gibt es Vorinformationen, die sich auf Erfahrungen in ähnlichen Prüfsituationen gründen oder die sich aus der Technologie des Produktionsprozesses ableiten lassen. Unter Verwendung eines Bayesschen Zuganges gelingt es, derartige Informationen geeignet quantifiziert in den Entscheidungsprozess mit einfließen zu lassen.

Ausgangspunkt ist hierbei, dass es möglich ist, unser Wissen vor dem Experiment über den unbekanntem Parameter p durch eine Wahrscheinlichkeitsverteilung mit der Dichte $f(p)$ auszudrücken. Diese Verteilung nennt man a-priori-Verteilung. Über die Bayessche Formel berechnet man dann aus der Stichprobenverteilung und der a-priori-Verteilung die so genannte a-posteriori-Verteilung. Diese enthält dann unser Wissen nach dem Experiment, jedoch unter Berücksichtigung von Vorinformationen und aktueller Stichprobeninformation (siehe z.B. Stephenson und Stern (1998) für eine Einführung in die Bayessche Methodologie).

Bayessche Formel:

$$f(p|k) = \frac{f(k|p)f(p)}{\int_{-\infty}^{+\infty} f(k|p)f(p)dp} \quad (11)$$

$f(k|p)$ Stichprobenverteilung, bei uns Binomialverteilung $B(n,p)$

$f(p)$ a-priori-Verteilung über den unbekanntem Parameter p

$f(p|k)$ a-posteriori-Verteilung über p

Im Zähler von (11) ergibt das Produkt von Stichprobenverteilung und a-priori-Verteilung gerade die gemeinsame Verteilung von k und p . Das Integral im Nenner ergibt die Randverteilung von k . Die gemeinsame Verteilung dividiert durch die Randverteilung von k ergibt die a-posteriori-Verteilung.

Günstig sowohl für die Modellierung wie für das Verständnis des Informationszuwachses ist es, wenn man zu gegebener Stichprobenverteilung eine a-priori-Verteilung findet, die zu einer a-posteriori-Verteilung vom selben Typ, wie die a-priori-Verteilung führt. Solche Verteilungspärchen nennt man konjugierte Verteilungsfamilien.

Wir wollen nun zeigen, dass die Betaverteilung konjugiert zur Binomialverteilung ist.

Mit der Stichprobenverteilung

$$f(k|p) = B(n, p) = \binom{n}{k} p^k (1-p)^{n-k}$$

und der a-priori-Verteilung

$$f(p) = \text{Beta}(\gamma, \delta) = \frac{\Gamma(\gamma + \delta)}{\Gamma(\gamma)\Gamma(\delta)} p^{\gamma-1} (1-p)^{\delta-1}$$

ergibt sich die gemeinsame Verteilung

$$\begin{aligned} f(k, p) &= f(k|p)f(p) = \binom{n}{k} p^k (1-p)^{n-k} \frac{\Gamma(\gamma + \delta)}{\Gamma(\gamma)\Gamma(\delta)} p^{\gamma-1} (1-p)^{\delta-1} \\ &= \binom{n}{k} \frac{\Gamma(\gamma + \delta)}{\Gamma(\gamma)\Gamma(\delta)} p^{k+\gamma-1} (1-p)^{n-k+\delta-1} \end{aligned} \quad (12)$$

Die Randverteilung $f(k)$ erhalten wir aus (12) durch Integration über p .

$$\begin{aligned} f(k) &= \int_0^1 \binom{n}{k} \frac{\Gamma(\gamma + \delta)}{\Gamma(\gamma)\Gamma(\delta)} p^{k+\gamma-1} (1-p)^{n-k+\delta-1} dp = \binom{n}{k} \frac{\Gamma(\gamma + \delta)}{\Gamma(\gamma)\Gamma(\delta)} \int_0^1 p^{k+\gamma-1} (1-p)^{n-k+\delta-1} dp \\ &= \binom{n}{k} \frac{\Gamma(\gamma + \delta)}{\Gamma(\gamma)\Gamma(\delta)} \frac{\Gamma(k + \gamma)\Gamma(n - k + \delta)}{\Gamma(n + \gamma + \delta)} \end{aligned} \quad (13)$$

Die a-posteriori-Verteilung ergibt sich nun als Quotient von gemeinsamer Verteilung (12) und Randverteilung (13) zu:

$$f(p|k) = \frac{f(k|p)f(p)}{f(k)} = \frac{\Gamma(n + \gamma + \delta)}{\Gamma(\gamma + k)\Gamma(\delta + n - k)} p^{\gamma+k-1} (1-p)^{\delta+n-k-1} \quad (14)$$

Die a-posteriori-Verteilung (14) ist aber wieder die Dichte einer Betaverteilung mit den neuen Parametern $\alpha = \gamma+k$ und $\beta = \delta+n-k$. Damit ist die Betaverteilung konjugiert zur Binomialverteilung.

Fassen wir zusammen:

$$\text{Beta}(\gamma, \delta) \text{ mit } B(n, p, k) \Rightarrow \text{Beta}(\gamma + k, \delta + n - k)$$

Verschmelzen wir die Vorinformation zu einer Ausfallrate aus einer Betaverteilung mit den Strukturparametern γ und δ mit dem Ergebnis einer binomialverteilten Stichprobe vom Umfang n , bei der insgesamt k Ausfälle auftreten, so erhalten wir als a-posteriori-Verteilung wieder eine Betaverteilung mit den neuen Strukturparametern $\gamma+k$ und $\delta+n-k$.

Parameter der Betaverteilung $Beta(\alpha, \beta)$:

$$\text{Dichtefunktion: } f(p) = \frac{\Gamma(\alpha + \beta)}{\Gamma(\alpha)\Gamma(\beta)} p^{\alpha-1} (1-p)^{\beta-1}, \quad \alpha, \beta > 0, \quad 0 \leq p \leq 1$$

Erwartungswert (Mittelwert) und Varianz einer betaverteilten Zufallsgröße p mit den Strukturparametern α und β ergeben sich zu:

$$E(p) = \frac{\alpha}{\alpha + \beta} \quad \text{Var}(p) = \frac{\alpha\beta}{(\alpha + \beta + 1)(\alpha + \beta)^2} \quad (15)$$

Der Modalwert (Maximum) der Betaverteilung liegt bei

$$\alpha > 1 \quad \text{und} \quad \beta > 1 \quad \Rightarrow \quad p_{\max} = \frac{\alpha - 1}{\alpha + \beta - 2}$$

$$\alpha = 1 \quad \text{und} \quad \beta > 1 \quad \Rightarrow \quad p_{\max} = 0$$

$$\alpha < 1 \quad \text{und} \quad \beta \geq 1 \quad \Rightarrow \quad p_{\max} = 0$$

$$\alpha = 1 \quad \text{und} \quad \beta = 1 \quad \Rightarrow \quad Beta(1,1) = 1 \quad (\text{constant})$$

$$\alpha > 1 \quad \text{und} \quad \beta = 1 \quad \Rightarrow \quad p_{\max} = 1$$

$$\alpha \geq 1 \quad \text{und} \quad \beta < 1 \quad \Rightarrow \quad p_{\max} = 1$$

$$\alpha < 1 \quad \text{und} \quad \beta < 1 \quad \Rightarrow \quad \text{Verteilung U-förmig(bimodal)}$$

Damit haben wir eine Vorstellung von der Gestaltvielfalt der Betaverteilung.

Verwenden wir als a-priori-Verteilung eine Betaverteilung mit den Parametern $\gamma = 1$ und $\delta = 1$, so ist das eine Gleichverteilung über dem Intervall $(0 ; 1)$. Wir sprechen hier von einer nichtinformativen a-priori-Verteilung, weil a priori jeder p -Wert im Intervall $(0 ; 1)$ die gleiche Chance bekommt.

Die Fiduzialverteilung (6) ergibt sich somit als a-posteriori-Verteilung in einem Bayes-schen Zugang bei nichtinformativer a-priori-Verteilung.

Konfidenzintervalle auf der Basis der a-posteriori-Verteilung bestimmt man so, dass die a-posteriori-Wahrscheinlichkeit über dem Intervall den gewünschten Schwellenwert (z.B. 0,95) nicht unterschreitet bei gleichzeitiger Minimierung der Intervallbreite. Solche Intervalle nennt man auch 95%-HPD-Intervalle (High Posterior Density).

Damit haben wir alle Begriffe und Zusammenhänge bereitgestellt, um unsere Prüfstrategie zu beschreiben.

Der folgende Abschnitt beschreibt, wie wir die Vorinformation in den Entscheidungsprozess einbinden.

Fallzahlbestimmung unter Einbeziehung von Vorinformationen

Die oben berechnete hohe Fallzahl von etwa 3.000 Proben würde sich also für den Fall ergeben, dass der Hersteller über keinerlei Informationen über die zu erwartende Ausfallrate verfügen würde (nichtinformativer Fall). Das ist jedoch hier nicht der Fall.

Einerseits erfolgt die Produktion der neuen Transplantate in einer hochsterilen Umgebung unter Verwendung von Ausgangsmaterial mit einem SAL von 10^{-6} , andererseits verfügt die Firma ABC GmbH über eine Stichprobe von bereits 1.046 auf Keimfreiheit geprüften Präparaten, die alle negativ getestet wurden.

Wir haben also zwei Informationsquellen:

1. Informationen über das verwendete Ausgangsmaterial, die Produktionsumgebung und das Herstellungsverfahren selber und
2. eine Stichprobe von 1.046 über längere Zeit auf Keimfreiheit geprüften Transplantaten, die in einem vergleichbaren Herstellungsverfahren unter vergleichbaren Bedingungen hergestellt wurden (siehe Anhang).

Wir schlagen deshalb folgende Teststrategie vor:

Wir starten mit einer Beta-a-priori-Verteilung mit den Parametern $\gamma = 1/3$ und $\delta = 1$. Dies ist keine Gleichverteilung, sondern trägt dem Umstand Rechnung, dass hier eine spezielle hochsterile Technologie zu Grunde liegt. Andererseits wird mit einer erwarteten Ausfallrate von 25% (ergibt sich aus (15)) bei dieser a-priori-Verteilung das Korsett nicht von Anfang an zu eng geschnürt. Die so angenommene a-priori-Verteilung dürfte das vorhandene Wissen über die erwartete Ausfallrate eher konservativ abschätzen. Es gibt jedoch viele Beta-a-priori-Verteilungen mit dem Erwartungswert von 0,25. Die hier (zunächst intuitiv) gewählte zeichnet sich unter allen derartigen Verteilungen dadurch aus, dass sie von allen unimodalen Betaverteilungen mit dem Erwartungswert von 0,25 diejenige ist, die am wenigsten informativ ist, die also bei festgehaltenem Erwartungswert die Varianz maximiert. Das wollen wir im Folgenden beweisen.

Für eine Wahl von $\gamma < 1$ und $\delta < 1$ ergeben sich U-förmige Dichtfunktionen, die als a-priori-Verteilungen hier keinen Sinn machen und deshalb ausgeschlossen werden.

Aus

$$E(p) = \frac{\gamma}{\gamma + \delta} = 0,25 \quad \text{folgt} \quad \gamma = \frac{\delta}{3}.$$

Aus der Formel (12) für die Varianz ergibt sich damit

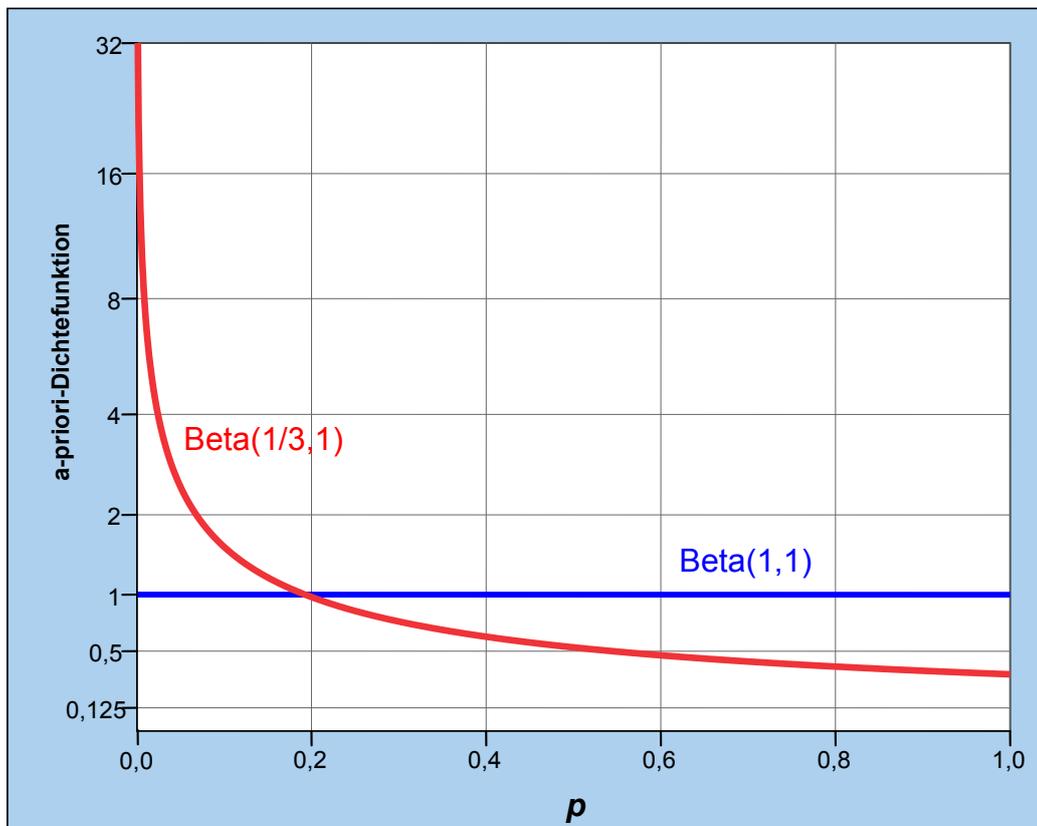
$$Var(p) = \frac{\delta^2 / 3}{(\delta/3 + \delta + 1)(\delta/3 + \delta)^2} = \frac{1/3}{\left(\frac{4}{3}\delta + 1\right)\frac{16}{9}} = \frac{3}{\frac{64}{3}\delta + 16} \Rightarrow \max_{\delta \geq 1}$$

Für $\delta = 1$ nimmt die Varianz ihr Maximum an: $Var(p) = 0,08$ oder $Std.Abw = 0,283$.

Die einzige Annahme in unserem Zugang, die nicht auf empirischen Daten basiert, ist die Wahl dieser a-priori-Verteilung, die wir hier einmal a-priori-Verteilung nullter Ordnung nennen wollen. Wüsste man am Anfang gar nichts, so würde man als a-priori-Verteilung nullter Ordnung eine nichtinformativ Gleichverteilung $Beta(1, 1)$ wählen.

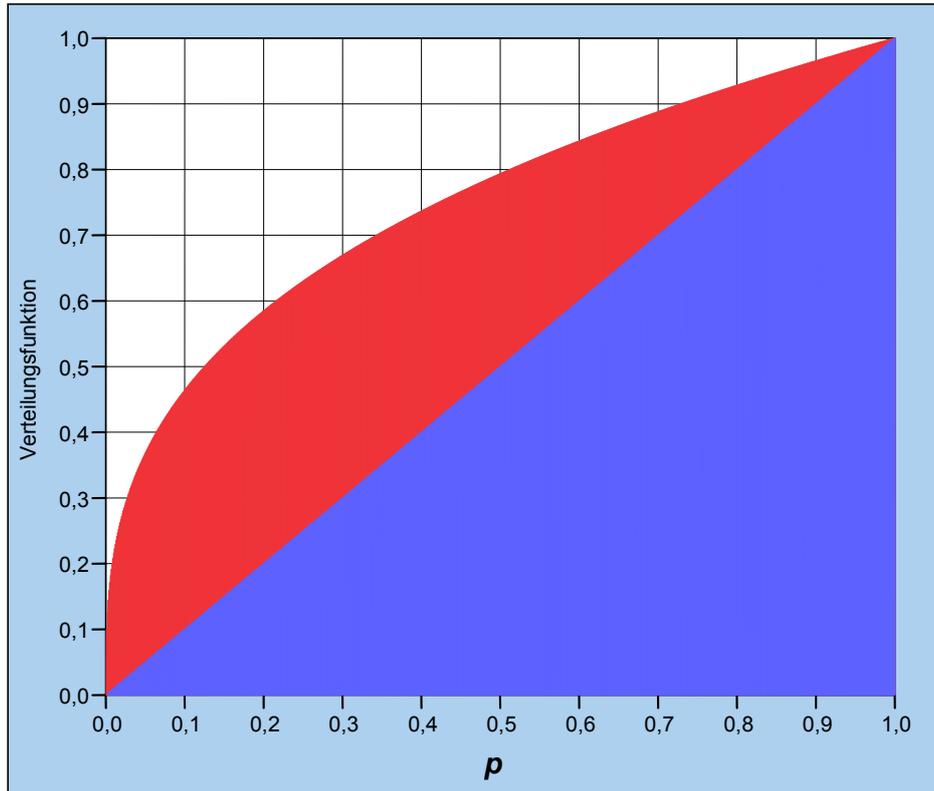
Die Herstellungstechnologie, die Ausgangsmaterialien und das erfahrene Personal rechtfertigen aber eine Abweichung von einer Gleichverteilung als a-priori-Verteilung, die für eine erwartete Ausfallrate von 50% steht. Wir haben deshalb die angenommene a-priori-Verteilung so geformt, dass diese im Rahmen der konjugierten Verteilungsfamilie liegt, also eine (sinnvollerweise unimodale) Betaverteilung ist, die halbe nichtinformativ Ausfallrate erreicht und unter diesen Nebenbedingungen am wenigsten informativ ist. So ergibt sich die hier vorgeschlagene a-priori-Verteilung in gewisser Weise ganz natürlich. Das einzige, was jetzt noch willkürlich ist, wäre der angenommene Erwartungswert von 0,25. Dieser besagt, dass bei allem, was wir über den Produktionsprozess des neuen Transplantates wissen, wir annehmen können, dass im Mittel nicht mehr als jedes 4. hergestellte Transplantat einen lebensfähigen Keim tragen wird. Oben wurde dazu schon argumentiert, dass diese Annahme sicher akzeptiert werden kann.

Informative $Beta(1/3,1)$ und nichtinformativ -a-priori-Verteilung



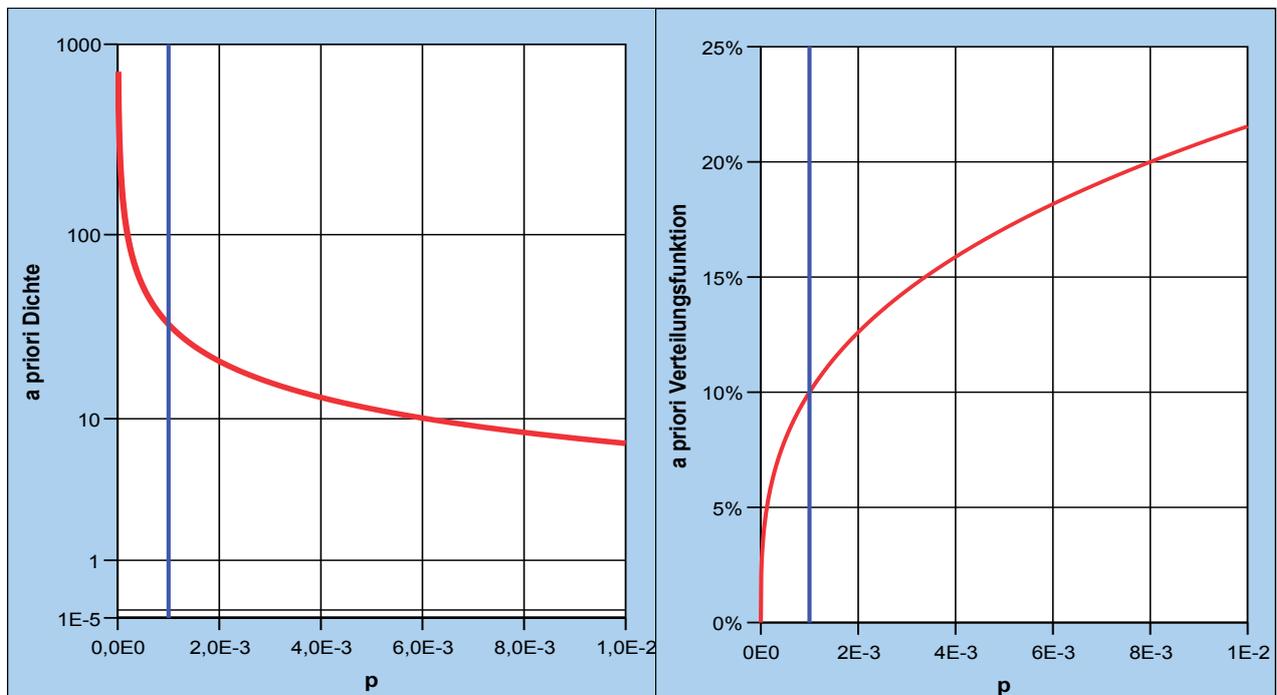
Zur besseren Darstellung wurde für die y-Achse hier eine logarithmische Skala gewählt.

a-priori- und nichtinformative Verteilungsfunktion



Die rote Fläche symbolisiert hier unser Zusatzwissen über die zu erwartende Ausfallwahrscheinlichkeit aus Herstellungstechnologie.

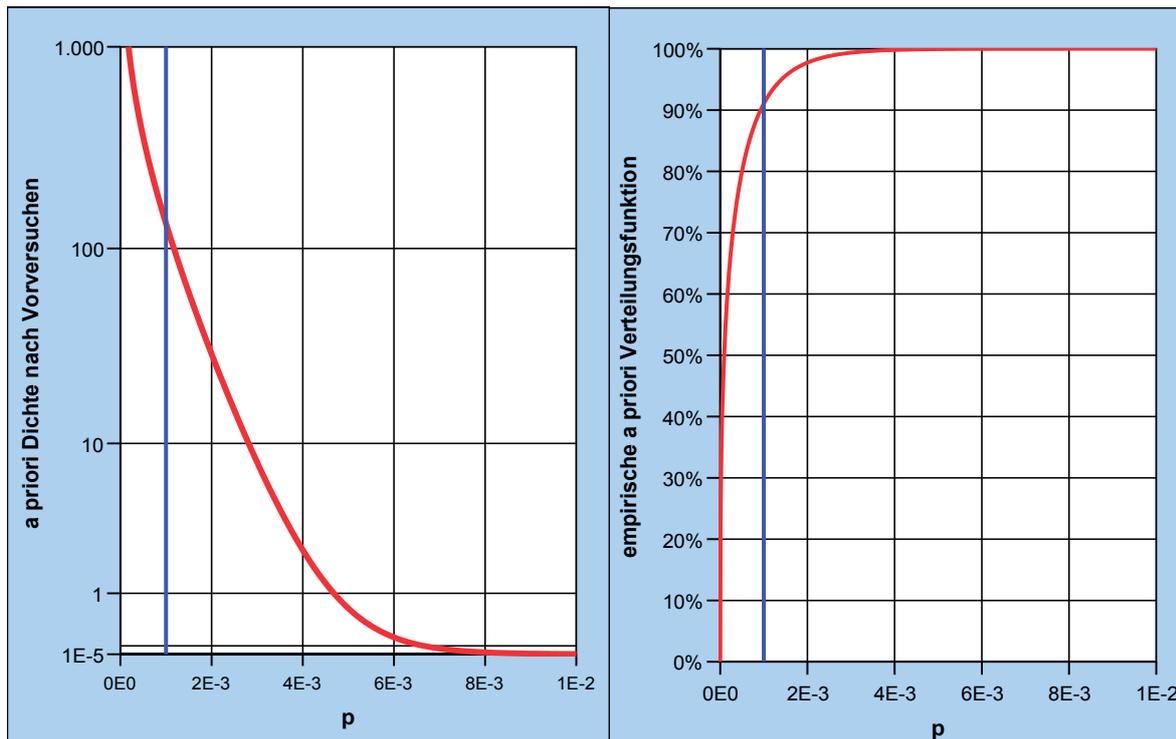
Beta-a-priori-Verteilung $\gamma=1/3$ und $\delta = 1$ (Ausschnitt)



Die a priori Wahrscheinlichkeit, dass $SAL \leq 0,001$, liegt bei unserer Konstruktion bei etwa 10%, also eine eher schwache Annahme.

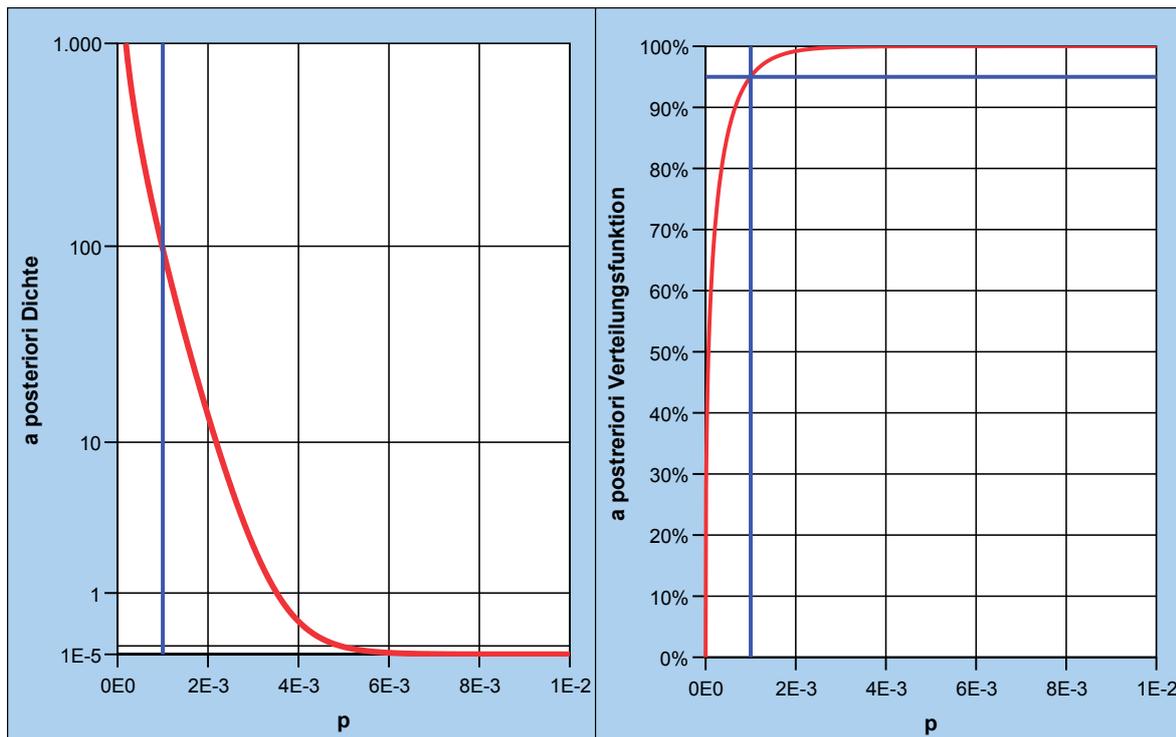
Diese a-priori-Verteilung verschmelzen wir jetzt mit den Informationen aus den in den letzten Jahren bei der ABC GmbH bereits vorgenommenen Produkttests. Die a-posteriori-Verteilung aus der Vorinformation und den 1.046 erfolgreichen Produkttests aus der Vergangenheit liefern unser neues a-priori-Wissen zum Beginn der neuen Testreihe, die zur Zulassung des neuen Transplantats führen soll.

A-priori-Verteilung nach 1.046 Vorversuchen



Nach 1.046 Vorversuchen (Erfahrungen bei früheren Tests) liegt die Wahrscheinlichkeit für $SAL \leq 0,001$ bei über 90%.

A-posteriori-Verteilung nach weiteren 432 keimfreien Proben



Nach der Totalprüfung von 3 Chargen á 144 keimfreien Transplantaten wird das geforderte $SAL \leq 0,001$ mit einer Sicherheit von 95% erreicht.

Die verwendeten Verteilungen:

Die **a-priori-Verteilung nullter Ordnung** (ohne empirische Vorinformationen):

$$f(p) = \text{Beta}(1/3, 1) = \frac{\Gamma(1,3333)}{\Gamma(0,3333)} p^{-0,6667}$$

Die **empirische a-priori-Verteilung** unter Verwendung früherer Testergebnisse:

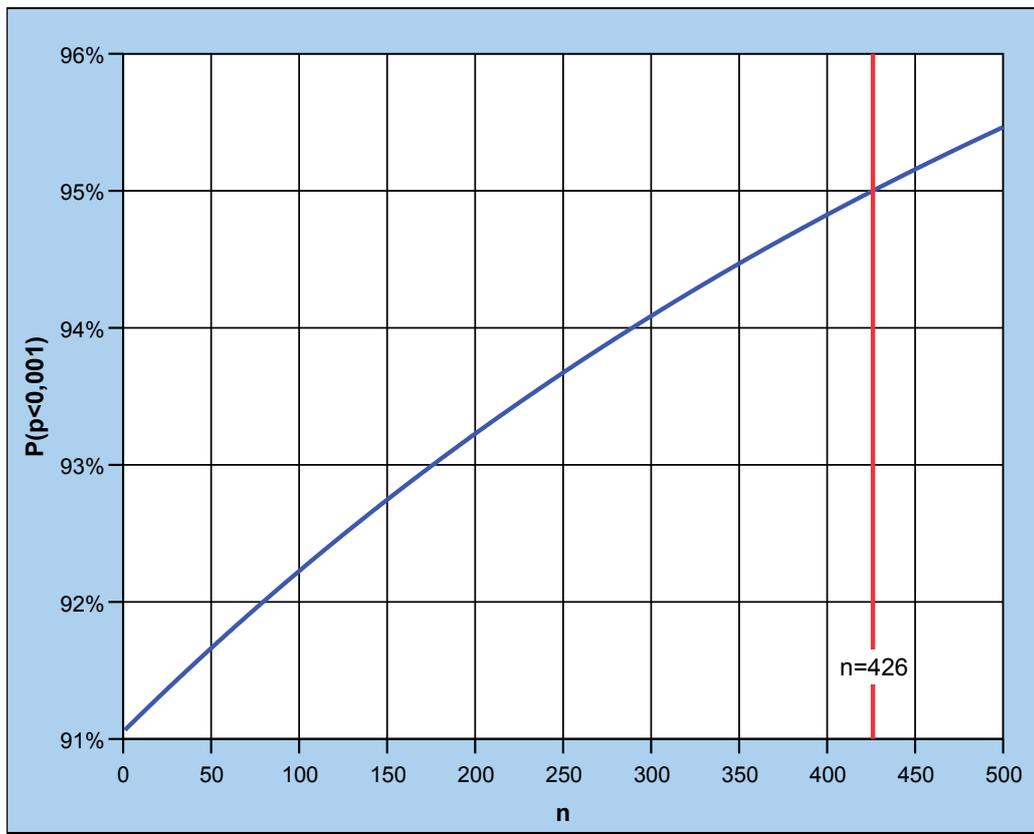
$$f(p | n_{\text{früher}}) = \text{Beta}(1/3, 1046 + 1) = \frac{\Gamma(1047,3333)}{\Gamma(0,3333)\Gamma(1047)} p^{-0,6667}(1-p)^{1046}$$

Die **a-posteriori-Verteilung** nach $n = 432$ keimfreien Proben:

$$f(p | n) = \text{Beta}(1/3, 1047 + 432) = \frac{\Gamma(1479,3333)}{\Gamma(0,3333)\Gamma(1479)} p^{-0,6667}(1-p)^{1478}$$

Das folgende Diagramm zeigt den Zusammenhang zwischen der aktuellen Fallzahl von keimfrei geprüften Proben und der erreichten Sicherheit für $SAL \leq 0,001$.

P(SAL ≤ 0,001) in Abhängigkeit von der Anzahl n der geprüften Proben



Nach $n = 426$ keimfreien Proben ist unter Einbeziehung aller verfügbarer Vorinformationen der geforderte Sterilitäts-Sicherheits-Schwellenwert (SAL) von $10^{\text{hoch}-3}$ mit einer Wahrscheinlichkeit von mindestens 95% erreicht.

Zusammenfassung

Mit der vorliegenden Expertise haben wir nachgewiesen, dass man bei einer attributiven statistischen Qualitätskontrolle bei seltenen Produktausfällen (hier ein lebensfähiger Keim im Transplantat) den Stichprobenumfang des Prüfplanes zur Erreichung des vorgeschriebenen Sterilitäts-Sicherheits-Schwellenwertes (SAL) von 10^{-3} unter Nutzung eines empirisch Bayesschen Zuganges deutlich reduzieren kann, wenn man vorhandene Vorinformationen über die unbekannte Ausfallrate geeignet quantifiziert im Entscheidungsprozess verwendet.

Für die Akzeptanz dieses Vorgehens sind **zwei Prämissen** wesentlich:

1. Man kann schon aus der Technologie des aseptischen Herstellungsprozesses Informationen über die zu erwartende Ausfallrate ableiten. Diese Information wird wie eine unscharfe Menge in Form einer Wahrscheinlichkeitsdichte modelliert. Als Funktionsklasse bieten sich die Betaverteilungen an, die neben einer großen Gestaltsvielfalt auch konjugiert zur Binomialverteilung ist. Es wurde versucht, hier eine akzeptanzfähige möglichst schwache Parametervorgabe zu treffen. Diese basiert auf der Überlegung, dass man bei Verwendung von keinerlei Vorinformationen die Gleichverteilung über dem Intervall $[0,1]$ als Repräsentant für unser unscharfes Wissen über die unbekannte Ausfallrate verwenden würde. Diese hat einen Erwartungswert (Mittelwert) von 0,5 (erwartete Ausfallrate also 50%). Unsere Grundannahme geht davon aus, dass die aseptische Technologie von vorn herein eine Reduktion der erwarteten Ausfallrate auf 25% zulässt. Aus der Menge aller univariaten Betaverteilungen wurde nun diejenige gewählt, die bei einem Erwartungswert von 0,25 am wenigsten informativ ist. Daraus ergibt sich die hier vorgeschlagene a-priori-Verteilung nullter Ordnung als $Beta(1/3,1)$.
2. Seit Januar 2005 hat die Firma ABC GmbH unter ganz ähnlichen Bedingungen bereits Transplantate in einer aseptischen Umgebung produziert und auf Keimfreiheit geprüft. Hier liegen empirische Daten vor, welche die aseptische Kompetenz des Antragstellers belegen. Diese Daten werden für die Validierung des neuen Herstellungsprozesses mit berücksichtigt. Aufbauend auf der a-priori-Verteilung nullter Ordnung wird – einem adaptiven empirisch Bayesschen Vorgehen folgend – mit jedem in der Vergangenheit getesteten Produkt (alle waren sie keimfrei) das Wissen über die unbekannte Ausfallrate immer präziser. Die a-posteriori-Verteilung nach allen Vorversuchen wird jetzt als a-priori-Verteilung für die neue Versuchsreihe verwendet.

Die Akzeptanz der Annahmen 1. und 2. sollte in diesem konkreten Fall möglich sein. Mit diesem Vorwissen gehen wir dann in die Planung des Stichprobenumfanges für die empirische Validierung des neuen aseptischen Herstellungsverfahrens. Nach 426 zerstörend geprüften Transplantaten, die sämtlich keimfrei sein müssen, ist dann die geforderte Sicherheit erreicht.

Noch eine Bemerkung am Ende:

Wenn eine attributive Qualitätskontrolle für einen laufenden Produktionsprozess nur durch eine Produktzerstörung möglich ist, ist es seit langem eine gängige Praxis aus jeder produzierten Charge eine kleinere Menge von Produkten (z. B. 20 von 144) zu entnehmen und diese zerstörend zu prüfen. Wenn diese keinen einzigen Keim enthalten, wird die Charge zum Verkauf freigegeben. Ein anderes Vorgehen ist wirtschaftlich gar nicht möglich. Auch hier wird der Erfahrung vertraut, auch diesem Vorgehen liegt implizit der Bayessche Ansatz zugrunde.

Wir haben diese stille Übereinkunft mathematisch formalisiert und hier auch auf die Phase der Produktneuzulassung angewendet. Ein Standardplan, der das tatsächlich vorliegende Wissen über die erwartete Ausfallrate ignoriert, würde etwa die Produktion und Zerstörung von etwa 3.000 Transplantaten erfordern, um das erforderliche SAL sicher nachzuweisen. Durch die Einbeziehung von vorhandenen Zusatzinformationen kann der Prüfplan auf $n = 426$ reduziert werden. Das spart Zeit, Geld und Ressourcen bei kontrollierter Sicherheit.

Literatur

- [1] Clopper, C.J., Pearson, E.S.: The use of confidence or fiducial limits illustrated in the case of the binomial. *Biometrics* **26**, 1934, 404-413.
- [2] Fisher, R.A.: The fiducial argument in statistical inference. *Annals of Eugenics*. **6**, 1935, 391 ff.
- [3] Gladitz, J. Fiduzialintervalle für den Parameter der Binomialverteilung mit SPSS 6.0 für Windows. *RZ-Mitteilungen der Humboldt-Universität* Nr. 9, Dezember 1994, 21-26.
- [4] Gladitz, J. Empirisch beste Tests bei zufälligem Parameter und ihre Anwendung in der Statistischen Qualitätskontrolle. *Math. Operationsforsch. u. Statist.* **7**, 1976, 685-696.
- [5] Stephenson, B., Stern, H. S.: A primer on the Bayesian approach to statistical inference. *Stats*, **23**, 1998, 3-9.
- [6] Zadeh, L. A., Fuzzy sets. *Information and Control*, Vol. 8, 1965, 338-353.

Berlin, im Januar 2007

Dr. Johannes Gladitz