Grundsätzlich unterscheidet man beim Hypothesentest, ob es sich um die **Bestätigung oder Verwerfung einer Hypothese** handelt, oder ob man **zwei Hypothesen gegenüberstellt**, um diejenige zu finden, die wahrscheinlicher ist.

Einfacher Hypothesentest

Gibt es nur eine Hypothese unterscheidet man zwischen einem linksseitigen und einem rechtsseitigen Hypothesentest.
Es kommt darauf an, Kriterien zu finden, nach denen die Hypothese

- abzulehnen ist oder
- nicht abgelehnt werden kann.

Grundsätzlich wird eine Hypothese nicht angenommen. Dass eine Hypothese nicht abgelehnt wird, heißt nicht automatisch, dass sie richtig ist, sondern nur, dass sie mit dieser Untersuchung nicht widerlegt werden kann. Andere Tests, die z.B noch nicht bekannt sind, könnten die Hypothese widerlegen.

Bei einem **linksseitigen Test** wird die Hypothese aufgestellt:

Liegt die Anzahl der Ereignisse **unter einer festgelegten Schranke**, dann ist die Hypothese abzulehnen. Alle Werte, die größer sind, widersprechen der Hypothese nicht.

Bei einem **rechtsseitigen Test** wird die Hypothese aufgestellt:

Liegt die Anzahl der Ereignisse **über einer festgelegten Schranke**, dann ist die Hypothese abzulehnen. Alle Werte, die kleiner sind, widersprechen der Hypothese nicht.

Bei einem **zweiseitigen Test** lautet die Hypothese:

Liegt die Anzahl der Ereignisse **unter einer Schranke** *oder* **über einer Schranke**, dann ist die Hypothese abzulehnen.
(Zumindest für das Abitur nicht relevant)

Die Anzahl der durchgeführten Proben bezeichnet man mit n, die Anzahl der positiv gefundenen Ereignisse mit k (entsprechend der Bedeutung in der Binomialverteilung).

Das Problem bei der Durchführung der Tests:
Auch Ereignisanzahlen, die über/unter der vorgegeben Schranke liegen sind durchaus möglich, aber wenn sie eintreten geht man **immer** davon aus, dass die Hypothese **falsch** ist, obwohl sie in wenigen Fällen auch richtig sein kann.

Eine Hypothese wird abgelehnt, obwohl sie richtig ist. In diesen Fällen spricht man von einem α – Fehler

Die Wahrscheinlichkeit für den Fehler ist genau die Summe der Wahrscheinlichkeiten, mit denen das Ereignis eintreten kann, bei dem die Anzahl k größer ist als die vorgegebene Grenze.

„Der Test verlief so ungünstig, dass er suggeriert, die Hypothese sei falsch, obwohl sie richtig ist.“
Dieser Fehler ist nicht zu erkennen, man muss mit ihm leben !

Für das Festlegen der Schranken für die Ablehnung/Annahme der Hypothese gibt es grundsätzlich zwei Möglichkeiten:

1. Ist die Anzahl k größer/kleiner als ein festgelegter Wert k_0, dann ist die Hypothese abzulehnen. Die Summe der Wahrscheinlichkeiten, mit denen trotzdem größere/kleinere Werte auftreten ist das Signifikanzniveau für den α – Fehler
2. Das Signifikanzniveau für den α – Fehler (Irrtumswahrscheinlichkeit) soll auf einen bestimmten Wert begrenzt werden. Ab/bis zu welcher Anzahl k_0 ist die Hypothese abzulehnen oder nicht abzulehnen.

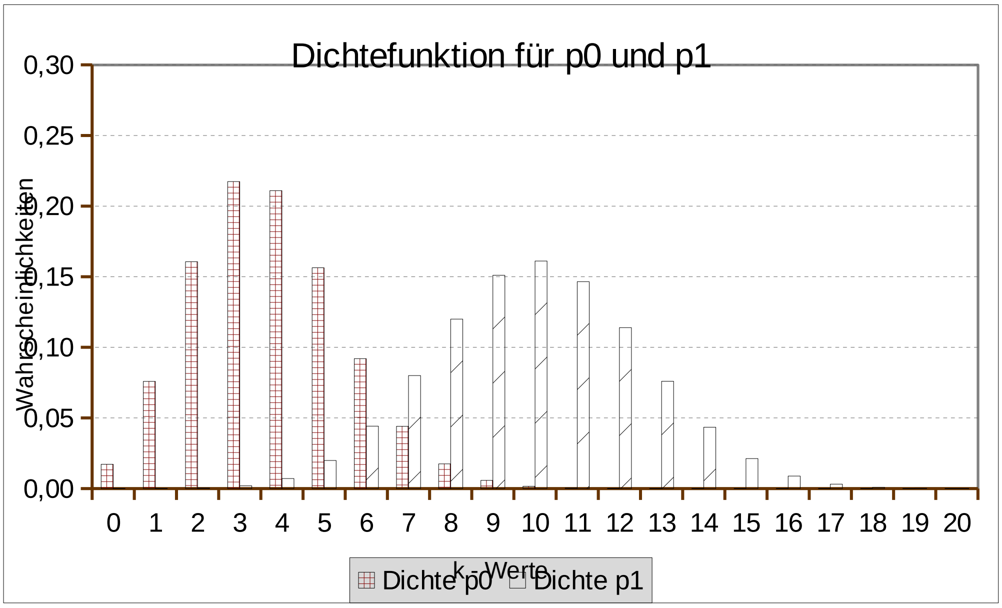
Alternativer Hypothesentest

Bei einem solchen Test wird die Aussage einer sogenannten Nullhypothese H_0 gegen eine alternative Hypothese H_1 getestet. Ein solcher Test führt immer dazu, dass es für die eine Hypothese ein linksseitiger Hypothesentest ist und für die andere Hypothese ein rechtsseitiger Hypothesentest ist.

Ein Hypothesentest für die Nullhypothese H_0 gegen die Alternative H_1 basierend auf den Beobachtungswerten x_1, x_2, ..., x_n besteht nun aus der Angabe eines kritischen Bereichs K (Verwerfungsbereich) zusammen mit der Entscheidungsregel:

- Verwerfe die Nullhypothese, falls die Beobachtungswerte im kritischen Bereich liegen.
- Belasse die Nullhypothese, falls nicht.

Ein statistischer Nachweis/Beweis gelingt bloß dann, wenn die Nullhypothese verworfen wird. Belassen der Nullhypothese bedeutet dagegen **nicht**, daß H_0 damit statistisch bewiesen ist!



Am besten lässt sich die Problematik an einem Bild veranschaulichen.
Die Nullhypothese entspricht den linken Balken im obigen Diagramm, die Alternativhypothese den rechten Balken. In beiden Fällen handelt es sich um die Dichtefunktion, dh. die Angabe der Wahrscheinlichkeit für jeweils einen Wert k (P(X=k)).
Die Nullhypothese gilt für k=0 bis k = 9 (oder10). Die Alternativhypothese von k=3 bist k = 17.
Was bedeutet das für den Test:
Werte von k > 10 werden wohl nicht auftreten, denn dann wäre die Summe der Wahrscheinlichkeiten bereits so nahe bei 1, dass man sie als sicher ansehen kann. Bei Werten k < 3 ist die Sache auch soweit klar, dass in diesem Bereich H_1 kaum auftreten kann. Was ist aber mit k = 6 oder k = 7. Ist diese Anzahl H_0 oder H_1 zuzuordnen.

Und an dieser Stelle kommt genau der α – Fehler zum Tragen:

Legt man die Grenze für k_0 zur Bestätigung oder Verwerfung von H_0 auf den Wert k= 6, dann gehören die Werte k > 6 zum α – Fehler. Man lehnt H_0 ab, obwohl es noch richtig sein könnte und entscheidet sich für H_1 . H_1 hat dort bereits eine höhere Wahrscheinlichkeit als H_0, trotzdem kann auch H_0 noch existieren.

Die Hypothese H_0 wird **abgelehnt, obwohl sie richtig ist**. In diesen Fällen spricht man von einem α – Fehler oder Fehler 1. Art

Was ist andererseits mit k = 5.

Bei k=5 glaubt man, sich im Bereich von H_0 zu befinden, weil dort H_0 eine höhere Wahrscheinlichkeit wie H_1 hat. Aber H_1 ist auch mit k = 5 und k = 4 möglich. In diesem Fall spricht man von einem β – Fehler, man nimmt H_0 an, obwohl H_1 richtig wäre.

Die Hypothese H_0 wird **angenommen, obwohl sie falsch ist**. In diesen Fällen spricht man von einem β – Fehler oder Fehler 2. Art

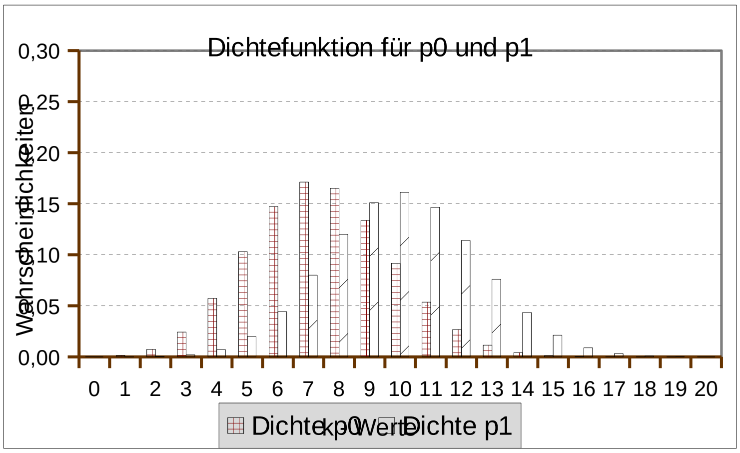
β – Fehler setzen immer eine zweite Hypothese voraus und sind für das Abitur nicht relevant.

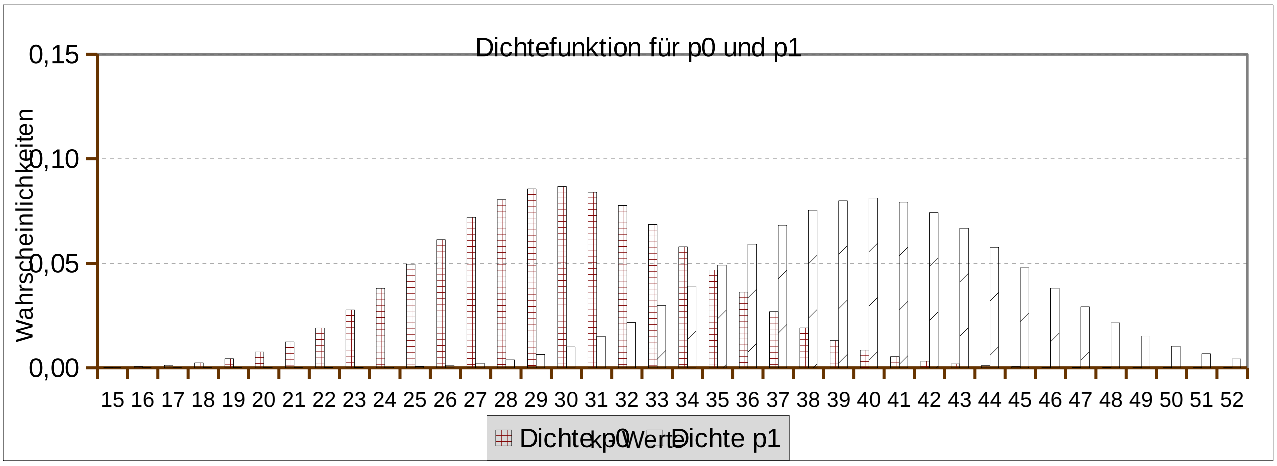
Die Nullhypothese und Alternative sind nicht gleichwertig:

- Wir wollen unbedingt vermeiden, daß wir die Nullhypothese verwerfen, obwohl sie wahr ist ("peinlicher Irrtum", Fehler 1. Art).
- Wenn unsere Untersuchung nicht genau genug ist (z.B. aufgrund einer zu geringen Stichprobengröße), kann es aber durchaus passieren, daß wir einen vorhandenen Effekt nicht
- entdecken, also die Nullhypothese belassen,obwohl sie falsch ist (Fehler 2. Art)

Bei einem Hypothesentest beschreibt das Signifikanzniveau die **(maximale) Wahrscheinlichkeit für den Fehler 1. Art**, die wir zulassen wollen. Für unseren Test soll also auf jeden Fall gelten:
P_{H0} [(X_1,X_2, ...,X_n) liegt im Verwerfungsbereich] ≤ α.
Das bedeutet für einen rechtsseitigen Test, dass die aufsummierte Wahrscheinlichkeit ab einem Wert k_0 kleiner als α sein muss

Aus den beiden Dichtfunktionen wird ein weiteres Problem sichtbar, dass man aus den Eigenschaften der Binomialverteilung ableiten kann. Liegen für ein Festes n die Wahrscheinlichkeiten dicht beieinander, dann sind die beiden Dichtefunktionen sehr ähnlich. Das rechte Bild zeigt die Dichtefunktionen für n=25 und p_0 =0,3 und p_1 = 0,4. Der Überlappungsbereich ist sehr groß, von k=4 bis k=14, bei k=8 und k= 9 sind die Wahrscheinlichkeiten annähernd gleich, aber sehr hoch. Damit wird kaum eine brauchbare Aussage möglich sein.





Das obere Bild zeigt die Dichtefunktionen für die gleichen Wahrscheinlichkeiten, aber für n = 100. Man erkennt deutlich, dass die beiden Dichtefunktionen weiter auseinandergezogen sind und die Wahrschein-lichkeiten an der Grenze (Balken von H_1 größer als der von H_0) bei 0,05 liegen, während sie bei n=25 bei 0,15 liegen, also wesentlich größer. Daraus ergeben sich zwei wesentliche Merkmale:

1. Je näher die Wahrscheinlichkeiten beieinander liegen, desto schlechter die Trennung in H0 und H1
2. Je kleiner der Stichprobenumfang, desto schlechter die Trennung in H0 und H1

Es soll hier auf die Festlegung der Trennstelle eingegangen werden. Bei welchem k legt man die kritische Zahl fest. Klar ist bei allen Fällen, wenn man den α – Fehler verkleinern will, erhöht man den β – Fehler und umgekehrt.
Je größer das k, desto kleiner der α –Fehler , und desto größer der β – Fehler
Je kleiner das k, desto kleiner der β – Fehler und desto größer der α –Fehler

Eine sinnvolle Trennstelle ist offenbar dasjenige k, an dem die Säule für die Hypothese H_0 kleiner wird, als die Säule für H_1. In diesem Beispiel ist das genau zwischen den Werten k = 34 und k = 35 der Fall.
Also ist einer dieser Werte eine gut geeignete Trennstelle.

Alternativtest bei Binomialverteilungen

Eine Fabrik liefert Schachteln mit Schrauben hoher Qualität (**10% der Schrauben sind fehlerhaft**) und minderer Qualität (**40% fehlerhaft**) an eine Baumarktkette.
Während des Ausladens geht bei einigen Verpackungen das Etikett ab. Da man nicht weiß, ob es sich um Schrauben 1. oder 2. Wahl handelt, muss ein Verfahren gefunden werden die Schrauben in kürzester Zeit der richtigen Qualität zuordnen zu können.
Weiter sei gegeben, dass sich in jeder Schachtel 300 Schrauben befinden.

Alternativtest: Man muss zwischen zwei Annahmen / Vermutungen / Hypothesen entscheiden.

(1) Formulieren der Hypothese

H0: "Die Schachtel ist hoher Qualität" Nullhypothese
H1: " Die Schachtel ist minderer Qualität" Alternative Hypothese

(2) Entscheidungsregel/ Prüfverfahren/ Test wird festgelegt

Da es zu zeitaufwendig wäre, alle 300 Schrauben zu untersuchen werden stattdessen nur **10 Schrauben** zufällig aus der Schachtel entnommen und getestet.
Dabei legt man fest, dass wenn **mehr als 3 Schrauben** unbrauchbar sind, die gesamte Schachtel der 2. Wahl angehören muss.

X := "Anzahl der schlechten Schrauben"

$0 \le x \le 3$: Die Schachtel ist gute Qualität
Entscheidung für H_0, gegen H_1 => Annahmebereich für H_0

$3 < x \le 10$: Die Schachtel ist minderer Qualität
Entscheidung für H_1, gegen H_0 => Annahmebereich von H_1, bzw.
Ablehnungs-/ kritischer Bereich von H_0

(3) Mögliche Fälle beim Prüfen

Realität	Entscheidung aufgrund des Testverfahrens	
H_0 ist wahr (p_0 = 0.1)	$0 \le x \le 3$ Annahme von H_0 richtige Entscheidung	$3 < x \le 10$ Ablehnung von H_1 falsche Entscheidung, **Fehler 1.Art**
H_1 ist wahr (p_1 = 0.4)	$0 \le x \le 3$ Ablehnung von H_0 falsche Entscheidung, **Fehler 2.Art**	$3 < x \le 10$ Annahme von H_1 richtige Entscheidung

Beim alternativen Test möchte man mit möglichst großer Wahrscheinlichkeit die richtige Entscheidung treffen. Trotzdem können falsche Zuordnungen vorkommen, jedoch im möglichst geringen Maß.

(4) Berechnung der Fehler 1.und 2. Art

Fehler 1. Art: **n = 10; p0 = 0,1; k > 3**

$P(X > 3) = 1 - F_{10;0,1}(3) = 0,01279$

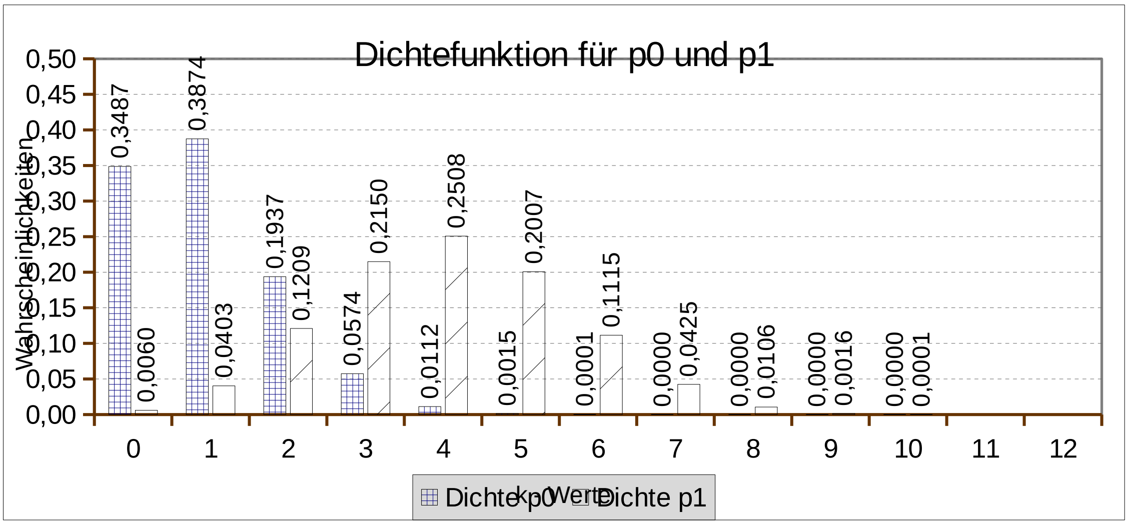
Die Wahrscheinlichkeit, dass die Kiste **gute Qualität** enthält (d.h. H_0 richtig ist), man sie aber dennoch **ablehnt** ist 1,28%. (Fehler 1. Art)

Fehler 2. Art: **n = 10; p1 = 0,4; k ≤ 3**

$P(X \le 3) = F_{10;0,4}(3) = 0,38228$

Die Wahrscheinlichkeit, dass die Kiste **2. Wahl** enthält (d.h. H_1 richtig ist), aber sie trotzdem **ablehnt** ist 38.23%.
Oder:
Die Wahrscheinlichkeit, dass **H_0 falsch** ist, man sie aber dennoch **als richtig annimmt** beträgt 38.23%.
Oder:
Irrtümlicherweise hält man 38.23% der schlechten Schachteln für gut.

n = 10
p0 = 0,1
p1 = 0,4



Die linken Säulen stellen die Dichtefunktion der Hypothese H_0 dar.
H_0 soll abgelehnt werden, wenn die Anzahl größer als 3 Stück beträgt.

Damit sind für den α – Fehler die Wahrscheinlichkeiten ab der Position k = 4 aufzuaddieren.

k	P(X=k)
4	0,01116
5	0,00149
6	0,00014

Summe: 0,01279

(s. Wert für den α – Fehler auf der vorhergehenden Seite)

Die rechten Säulen stellen die Dichtefunktion der Hypothese H_1 dar.
H_1 soll abgelehnt werden, wenn die Anzahl weniger als 4 Stück beträgt.

Damit sind für den β – Fehler die Wahrscheinlichkeiten bis zur Position k = 3 aufzuaddieren.

k	P(X=k)
0	0,0060
1	0,0403
2	0,1209
3	0,2150

Summe: 0,3822

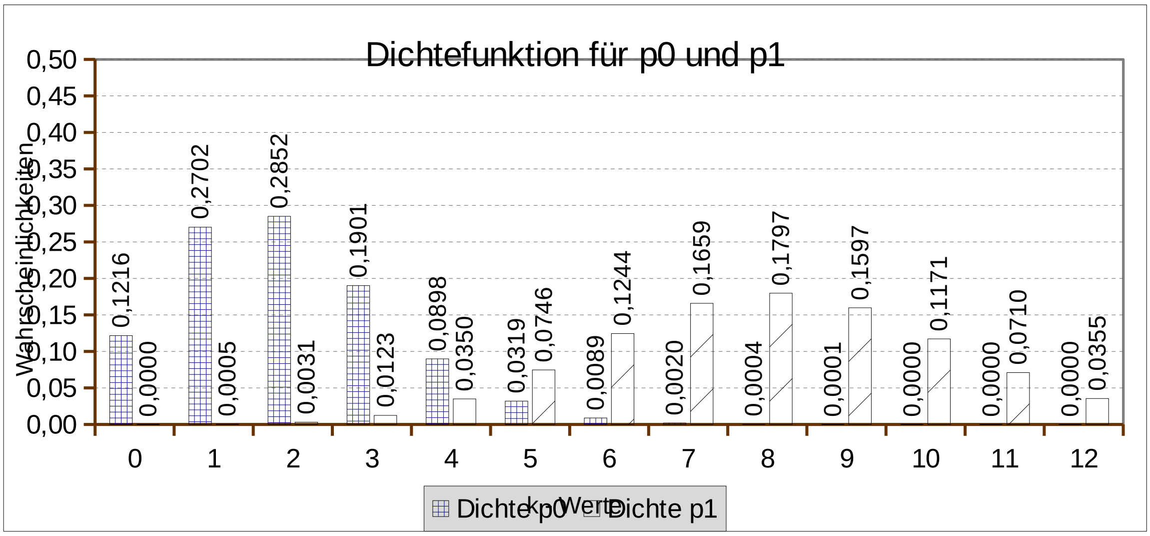
(s. Wert für den β – Fehler auf der vorhergehenden Seite)

Die Wahrscheinlichkeit für den Fehler ist dem Baumarkt zu groß. Das ist die Wahrscheinlichkeit, dass eine Schraubenkiste nur 2. Qualität hat, aber als 1. Qualität eingestuft wird. Deshalb will man den Test verändern. Da die Stichprobengröße mit 10 sehr klein gehalten wurde, entscheidet man sich die Stichprobengröße zu verändern.

Die Stichprobengröße wird auf 20 Schrauben erhöht, die Kiste wird als 2. Qualität eingestuft, wenn mehr als 4 schlechte Schrauben dabei sind.

Das Problem soll an Hand der Funktionsbilder untersucht werden.

n = 20
p0 = 0,1
p1 = 0,4



Der α – Fehler liegt jetzt bei 0,04317
und der β – Fehler bei 0,05095
Auf die Berechnung soll hier nicht noch einmal eingegangen werden.

Linksseitiger Hypothesentest bei Binomialverteilungen

Ablaufschema beim linksseitigen Hypothesentest

1) Festlegung der Hypothesen: H_0: $p \geq p_0$; (H_1: $p < p_0$)

   (beim linksseitigen Hypothesentest immer **$p \geq p_0$**)
2) Festlegung des Stichprobenumfangs n und der Irrtumswahrscheinlichkeit α
3) Festlegung der Zufallsvariablen X (immer) und ihrer Verteilung (verlangen manche Lehrer nicht; in der Schule ist es fast immer die Binomialverteilung)
4) Bestimmung der linken Grenze gL aus der Bedingung $P(X \leq g_L) \leq \alpha$ und damit des Ablehnungsbereichs K = {0, …, g_L}
5) Angabe der Entscheidungsregel oder Entscheidung aufgrund eines konkreten Stichprobenergebnisses

Bei einem linksseitigen Hypothesentest sprechen kleine Werte der Zufallsvariablen gegen die Hypothese, also Werte, die links auf dem Zahlenstrahl bzw. links vom Erwartungswert liegen. Wenn eine Partei behauptet, sie habe einen Stimmenanteil von mindestens 40 %, macht es uns stutzig, wenn zu wenige Leute behaupten, diese Partei wählen zu wollen.

Beispiel

Bei einer Wahl hatte eine Partei einen Stimmenanteil von 40 %. Nach der Wahl hat sie einige unbequeme Maßnahmen ergriffen, und man vermutet, dass der Stimmenanteil gesunken ist. Bei einer Umfrage unter 100 Personen geben 33 an, dass sie die Partei wieder wählen würden. Kann man hieraus bei einer Irrtumswahrscheinlichkeit von 5 % schließen, dass der Stimmenanteil gesunken ist?

Lösung

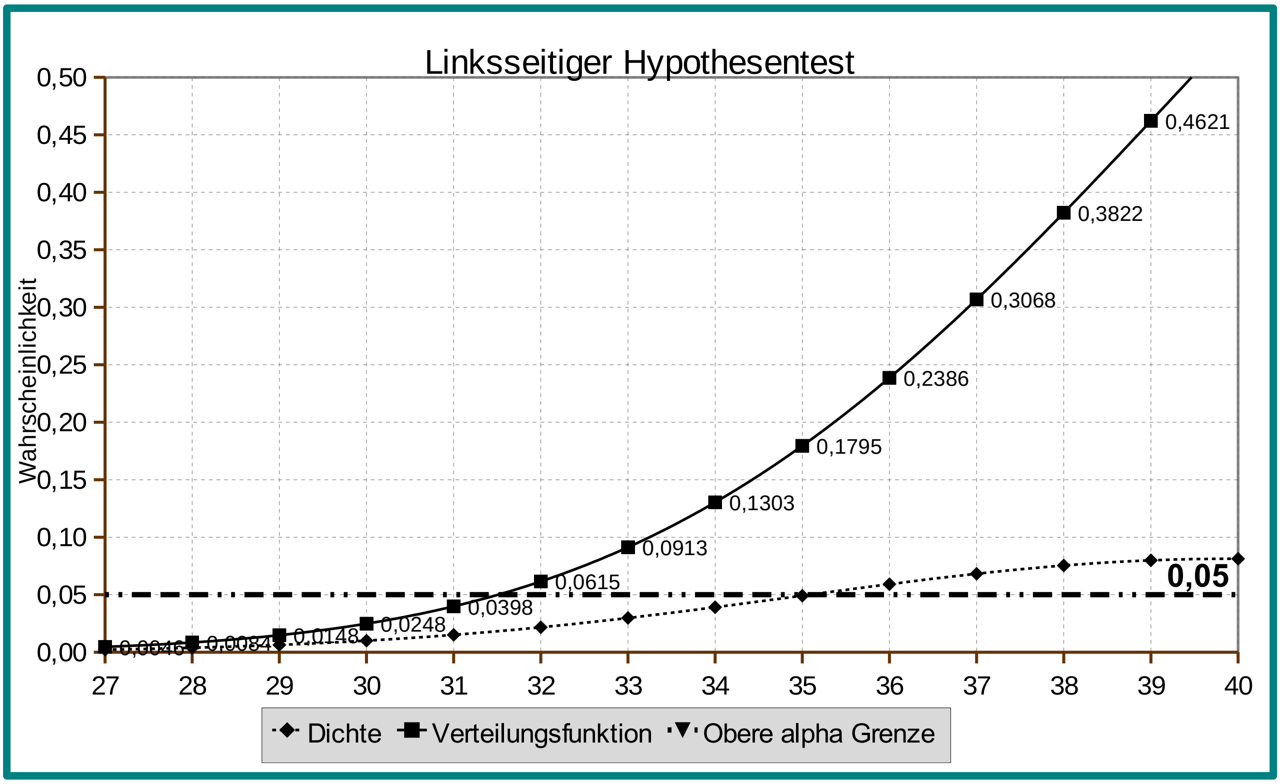
1) Wir führen den Hypothesentest nach dem Muster durch:

   H_0: $p \geq 0{,}4$; H_1: $p < 0{,}4$
2) $n = 100$; $\alpha = 0{,}05$
3) X = Anzahl der Personen in der Stichprobe, die die Partei wieder wählen würden

   X ist bei wahrer Nullhypothese binomialverteilt mit den Parametern n = 100 und p = 0,4
4) $P(X \leq g_L) \leq 0{,}05 \Rightarrow g_L = 31$ (Tabelle siehe Rückseite); Ablehnungsbereich K = {0, …, 31}
5) Da $33 \notin K$ ist, kann man bei einer Irrtumswahrscheinlichkeit von 5 % nicht behaupten, dass der Stimmenanteil der Partei gesunken ist.

   Wäre kein Stichprobenergebnis bekannt, so würde man an dieser Stelle die Entscheidungsregel formulieren: Wenn höchstens 31 Leute angeben, die Partei wieder wählen zu wollen, kann man bei einer Irrtumswahrscheinlichkeit von 5 % behaupten, dass der Stimmenanteil der Partei gesunken ist.

Wertetabelle für Summierte
Binomialverteilung für n = 100
und p = 0,4



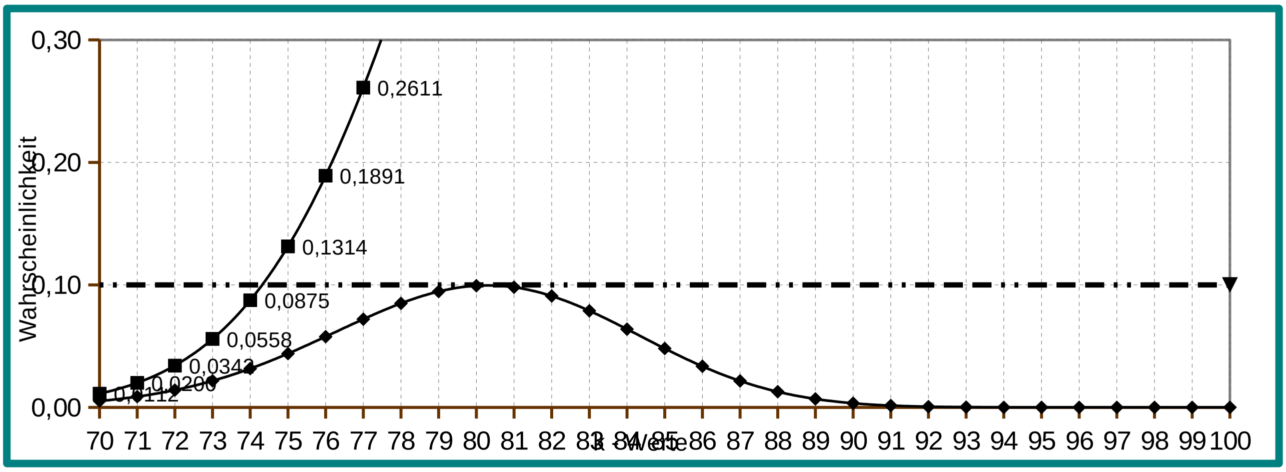
k	$B_{100;0,4}(k)$	$F_{100;0,4}(k)$
27	0,00220	0,004600
28	0,00383	0,008433
29	0,00634	0,014775
30	0,01001	0,024783
31	0,01507	0,039848
32	0,02166	0,061504
33	0,02975	0,091254
34	0,03908	0,130337
35	0,04913	0,179469
36	0,05914	0,238611
37	0,06820	0,306810
38	0,07538	0,382188
39	0,07989	0,462075
40	0,08122	0,543294

Die strich-punktierte Linie kennzeichnet die Grenze für die Irrtums-wahrscheinlichkeit von 5% = 0,05. Dieser Wert liegt zwischen den k-Werten 31 und 32.
Wenn sich immer noch 33 Personen für diese Partei entscheiden, dann liegt dieser Wert nicht im Ablehnungsbereich.
Man kann daraus nicht schließen, dass der Stimmenanteil gesunken ist.

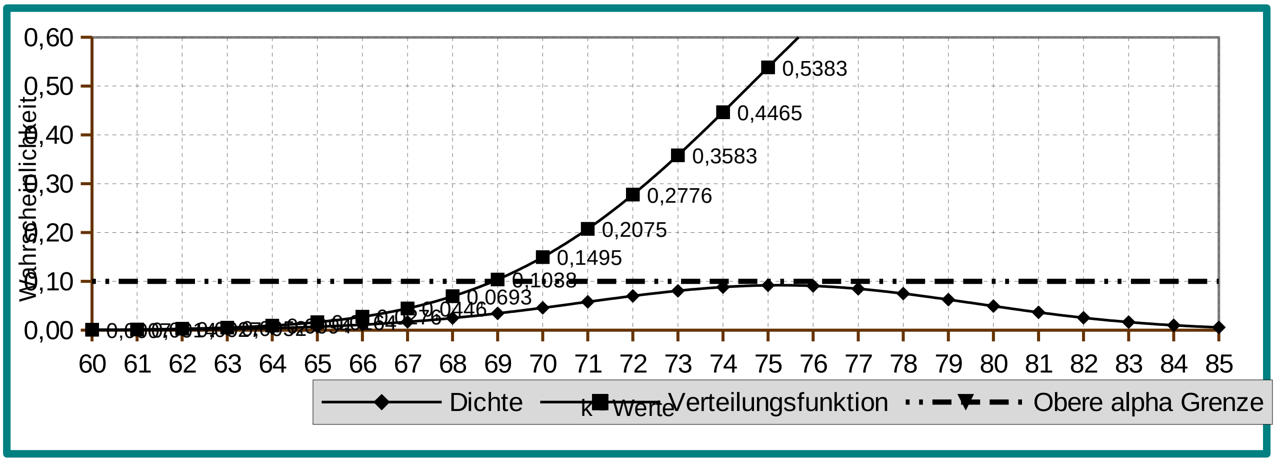
Warum steht beim linksseitigen Hypothesentest: H_0 : $p \geq p_0$

Der Anbieter eines Medikamentes sagt: Sein Produkt hat eine Heilungsquote von 8% (p_0 = 8%). Wir als Nutzer haben Zweifel an der Quote und meinen er übertreibt : p < 8%. **Das ist die Hypothese H_1.** Wenn die Heilungs-quote größer ist, haben wir Pech gehabt, also gehören alle Wahrscheinlichkeiten $p > p_0$ zur positiven Seite des Anbieters, damit gehören zur Anbieterseite: $p \geq p_0$. **Das ist die Hypothese H_0**, die wir widerlegen wollen, weil unsere Auffassung lautet: p < 8%. Das bedeutet, wenn bei einer Auswahl von Personen sehr wenige Personen geheilt sind, dann spricht das gegen die Aussage des Anbieters. Damit liegt der Ablehnungsbereich links, bei kleinen Werten, damit ein Linksseitiger Hypothesentest.

H_0: **$p \geq 0{,}8$** und n = 100 und als Signifikanzniveau α = 0,1



Die Grenze des Ablehnungsbereichs für p = 0,8 liegt bei k = 74. Für p = 0,8 tritt der Wert k = 74 selbst nur mit einer Wahrscheinlichkeit von 3,1 % auf, der Wert 73 mit einer Wahrscheinlichkeit von 2,1 %.



Für p = 0,75 treten Werte von k < 74 schon mit einer Wahrscheinlichkeit von 44,6% auf, dh. sie sind viel häufiger. Der Wert von k = 74 allein tritt mit einer Wahrscheinlichkeit von 8,8%. Erscheinen also Werte, die kleiner als 74 sind, dann spricht das eher für eine kleiner Wahrscheinlichkeit, da der Wert bei 8% doch relativ selten auftritt.

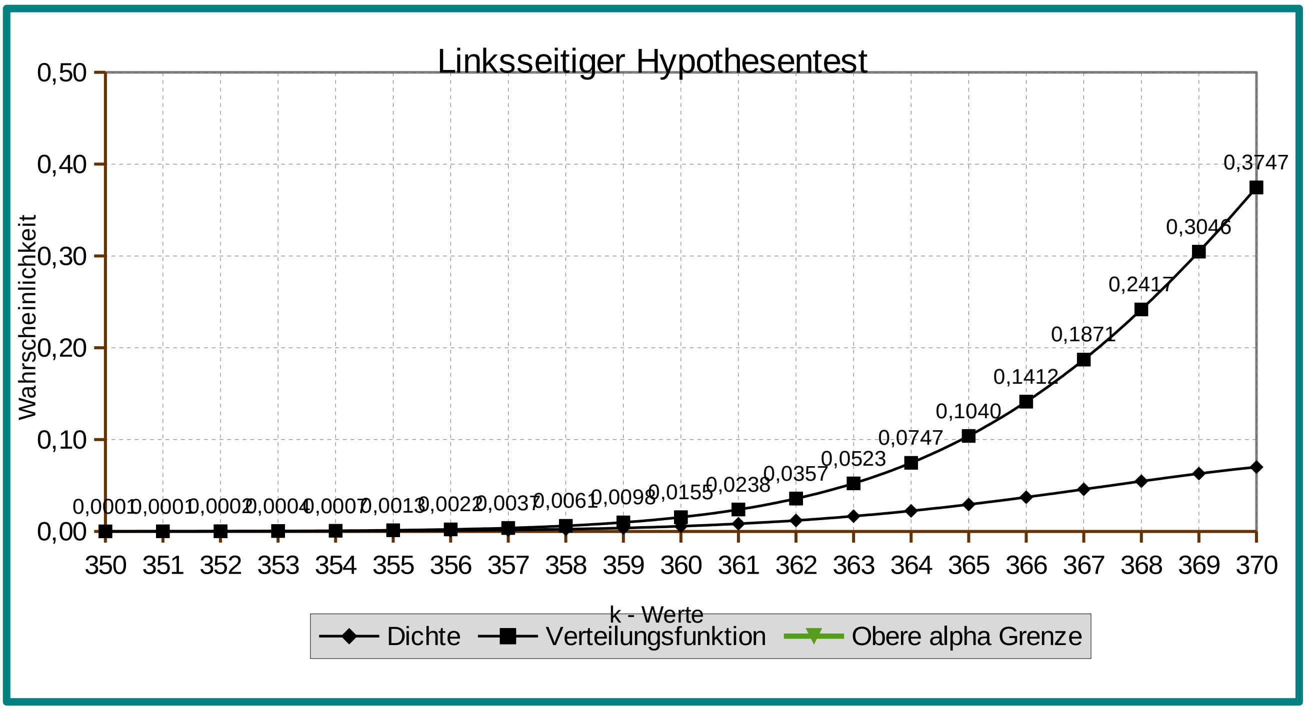
Mit der **kleinsten angegebenen Wahrscheinlichkeit** ist der

- Ablehnungsbereich bereits so weit nach unten verschoben, dass man bei noch kleineren Werten die angegebene Wahrscheinlichkeit nicht mehr glauben kann.
- der Annahmebereich vom Erwartungswert aus so weit nach links verschoben ist, dass die Möglichkeit eines Auftretens noch kleinerer k Werte als selten bei der angegebenen Wahrscheinlichkeit ansehen muss.

Im angegebenen Fall liegt der Erwartungswert bei E(X) = n • p = 100 • 0,08 = 80. Man räumt die Möglichkeit ein, dass man zufällig eine ungünstige Probe zieht und lässt noch zu, dass weniger Personen von dem Medikament geheilt werden, ab noch weniger als 74 lassen Zweifel aufkommen.

Beispiel.
Ein Großhändler verkauft Glühbirnen an einen Baumarkt. Er gibt an, dass von den Glühbirnen (mindestens) **95%** in Ordnung sind. Der Baumarkt zieht eine Probe von **400 Glühbirnen** und will auf Grund dieser Stichprobe entscheiden, ob er eine größere Menge Glühbirnen bestellt. Wie könnte eine Entscheidungs-regel lauten, so dass der Baumarkt in **95%** der Fälle richtig liegt?

Die Art des Tests
Da man durch einen Test eine Hypothese nie beweisen kann, sondern sie bei einer großen Abweichung höchstens verwerfen kann, wird der Baumarkt für den Test davon ausgehen, dass weniger als 93% der Glühbirnen in Ordnung sind. In diesem Fall p_0 = 93%. Als Abnehmer stellen wir uns auf den Standpunkt, die Angabe des Anbieters ist falsch (Er „lügt“). Dann ist für uns als Abnehmer der Fall uninteressant, wenn mehr als 93% funktionierende Glühbirnen vorhanden sind, für uns ist nur interessant, wenn weniger als 93% intakte Glühbirnen vorhanden sind. Damit ist unsere Behauptung $p < 93\%$. **Das ist die Hypothese H_1 !**
Damit bleibt für die Aussage des Anbieters als Erweiterung $p \geq 93\%$. **Das ist die Hypothese H_0**, der wir nicht trauen und die wir widerlegen wollen.
Kleinere Wahrscheinlichkeiten bringen eine Binomialverteilung hervor die sich nach links verschiebt. Damit verschiebt sich der Erwartungswert (höchste Säule) und seine Umgebung (die Werte, die am häufigsten auftreten) auch nach links. Das stört uns nicht.
Größere Wahrscheinlichkeiten verschieben die Funktion nach rechts. Das heißt, kleinere Wahrscheinlich-keiten verschieben den Erwartungswert nach links. Das wollen wir nicht, das ist für uns ungünstig. Also sagen wir: Kleine Werte k, die kleiner, als ein von uns festgelegtes k_0 sind, dürfen nur selten auftreten. Damit sollen Werte, die kleiner als k_0 sind insgesamt (alle noch kleineren) höchstens mit einer Wahrscheinlichkeit von α (Signifikanzniveau) auftreten. Damit entsteht ein linksseitiger Test, da der Ablehnungsbereich am linken Ende der Wahrscheinlichkeitsfunktion auftritt.



Interpretation der Kurve:
Bei k_0 = 362 wird das erste mal die Grenze von 5% unterschritten. Findet man bei den 400 Glühbirnen weniger als 362 intakte Glühbirnen, sind Zweifel an den 93 % angebracht. Der Erwartungswert ist in diesem Fall
$E(X) = n \cdot p = 400 \cdot 0{,}93 = 372$. Bei einer „idealen“ Stichprobe sollte man 372 intakte Glühbirnen finden. Natürlich kann man auch weniger finden, wenn man ungünstig zieht. Aber weniger als 362 ist schon unwahrscheinlich. k = 362 allein tritt schon nur noch mit einer Wahrscheinlichkeit von 1,2% und alle kleineren Ergebnisse noch seltener. Da sind Zweifel an der Wahrscheinlichkeit von 93% gegeben. Geringere Wahrscheinlichkeiten verschieben die Kurven nach links, damit treten kleiner Werte häufiger auf.

Rechtsseitiger Hypothesentest bei Binomialverteilungen

Ablaufschema beim rechtsseitigen Hypothesentest

1) Festlegung der Hypothesen: H_0: $p \leq p_0$; (H_1: $p > p_0$)

   (beim rechtsseitigen Hypothesentest immer $\mathbf{p \leq p_0}$)
2) Festlegung des Stichprobenumfangs n und der Irrtumswahrscheinlichkeit α
3) Festlegung der Zufallsvariablen X (immer) und ihrer Verteilung (verlangen manche Lehrer nicht; in der Schule ist es fast immer die Binomialverteilung)
4) Bestimmung der rechten Grenze g_R aus der Bedingung $P(X \geq g_R) \leq \alpha$ und damit des Ablehnungsbereichs K $= \{g_R, \ldots, n\}$
5) Angabe der Entscheidungsregel oder Entscheidung aufgrund eines konkreten Stichprobenergebnisses

Bei einem rechtsseitigen Hypothesentest sprechen große Werte der Zufallsvariablen gegen die Hypothese, also Werte, die rechts auf dem Zahlenstrahl bzw. rechts vom Erwartungswert liegen. Wenn eine Firma behauptet, sie habe bei der Produktion von Taschenrechnern eine Ausschussquote von höchstens 5 %, macht es uns stutzig, wenn zu viele defekte Taschenrechner reklamiert werden.

Beispiel

Die Behauptung der Firma soll mit einer Stichprobe von 100 Taschenrechnern untersucht werden. Wie viele defekte Taschenrechner müssen mindestens gefunden werden, damit man bei einer Irrtumswahrscheinlichkeit von 10 % schließen kann, dass die Ausschussquote höher als von der Firma angegeben ist?

Lösung

Wir führen den Hypothesentest nach dem Muster durch:

1) H_0: $p \leq 0{,}05$; H_1: $p > 0{,}05$
2) $n = 100$; $\alpha = 0{,}10$
3) X = Anzahl der defekten Taschenrechner in der Stichprobe

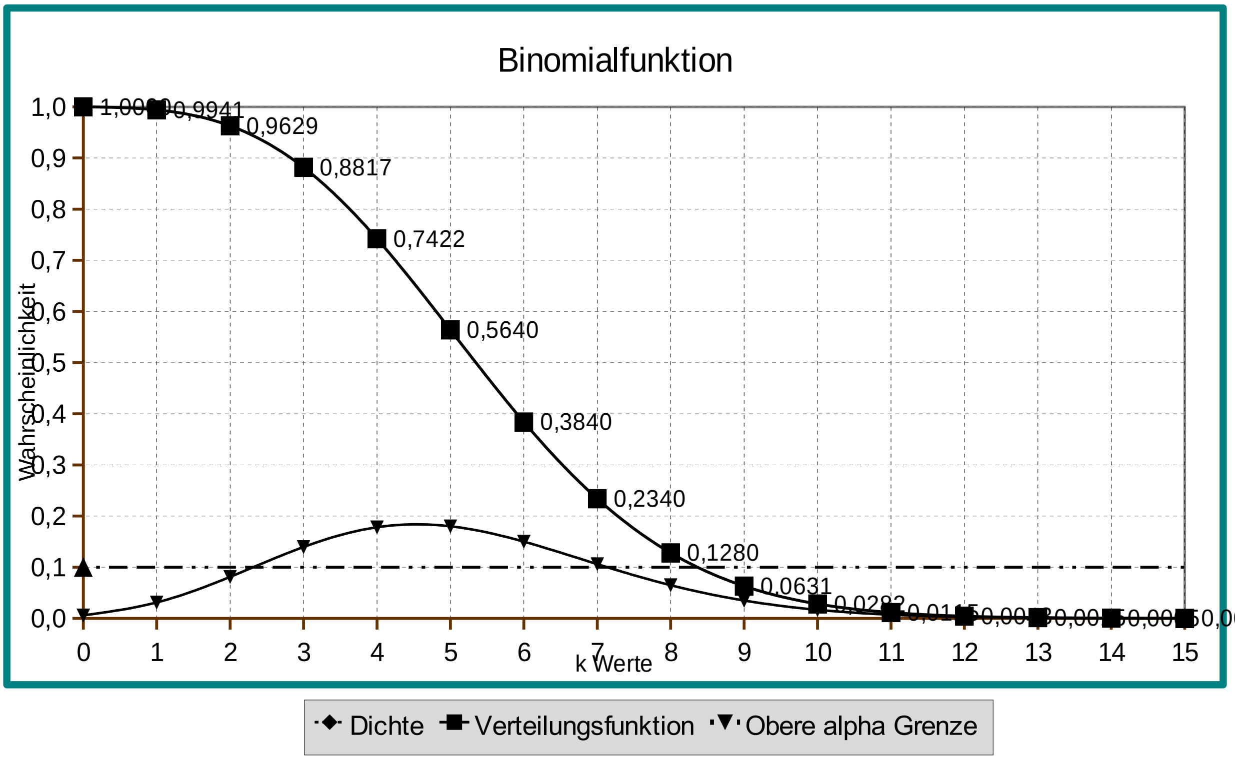
   X ist bei wahrer Nullhypothese binomialverteilt mit den Parametern n = 100 und p = 0,05
4) Ansatz: $P(X \geq g_R) \leq 0{,}10$
5) Da eine solche Wahrscheinlichkeit nicht direkt aus der Tabelle abgelesen werden kann, muss man zunächst umformen bzw. zum Gegenereignis übergehen:

   $P(X \geq g_R) \leq 0{,}10 \Leftrightarrow P(X \leq g_{R-1}) \geq 0{,}90 \Rightarrow g_{R-1} = 8$ (aus Tabelle) $\Leftrightarrow g_R = 9$;

   Ablehnungsbereich $K = \{9, \ldots, 100\}$
6) Wenn mindestens 9 defekte Taschenrechner entdeckt werden, kann man bei einer Irrtumswahrscheinlichkeit von 10 % behaupten, dass die Angaben der Firma nicht zutreffen.

   Nur wenn ein Stichprobenergebnis bekannt wäre, könnte man ein konkretes Urteil fällen.

Wertetabelle für die umgekehrte
Summierte Binomialverteilung für
n = 100 und p = 0,05



	Dichte	umgekehrte Verteilungsfunktion
0	0,005920	1,000000
1	0,031161	0,994079
2	0,081182	0,962919
3	0,139576	0,881737
4	0,178143	0,742161
5	0,180018	0,564019
6	0,150015	0,384001
7	0,106026	0,233986
8	0,064871	0,127960
9	0,034901	0,063090
10	0,016716	0,028188
11	0,007198	0,011472
12	0,002810	0,004274
13	0,001001	0,001464
14	0,000327	0,000463
15	0,000099	0,000136

An der umgekehrten Verteilungsfunktion (Linie mit den Quadraten) ist zu erkennen, dass die Irrtumswahrschein-lichkeit von 10% = 0,1 zwischen den k Werten8 und 9 liegt.
Bei k = 9 kann man bei einer Irrtumswahrscheinlichkeit von 0,1 davon ausgehen, dass die Fehlerrate von 5% nicht stimmt.
Die Summe von 9 bis 30 ergibt 0,0631 (s. Kurvenbild). Die Summe von 8 bis 30 ergibt 0,1280, also mehr als 0,1

Der Erwartungswert E(X) ist für diese Binomialverteilung
E(X) = n • p = 100 • 0,05 = 5
und die Streuung
$\delta = \sqrt{n\ p\ q} = \sqrt{100 \cdot 0{,}05 \cdot 0{,}95} = 2{,}18$

ACHTUNG   Bei der Berechnung der umgekehrten Verteilungsfunktion

Die summierte Verteilungsfunktion binomcdf(n,p,k) liefert für n = k immer den Wert 1, da dann alle möglichen Wahrscheinlichkeiten summiert sind. Damit wird bei der Berechnung von 1 – binomcdf(n,p,n) immer der Wert 0 entstehen. Aber die Einzelwahrscheinlichkeit für n ist niemals 0, sondern immer ein Wert größer 0, nämlich binompdf(n,p,n). Deshalb ist bei der umgekehrten Wahrscheinlichkeit immer ein k Wert weniger zu benutzen. Damit sieht die notwendige Formel so aus:

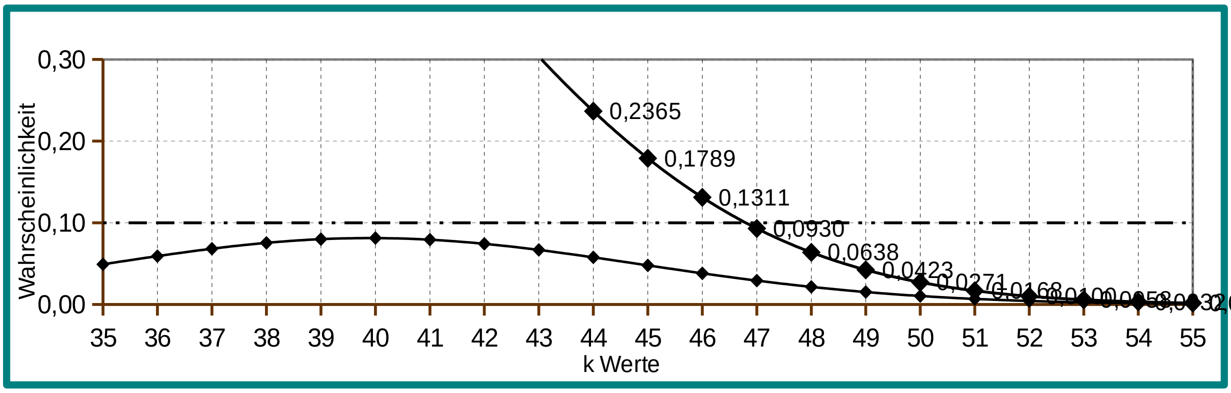
1 – binomcdf(n,p,**k-1**)

Die Werte von k sind bei der Eingabe in den GTR jeweils die X-Werte in der ersten Spalte.

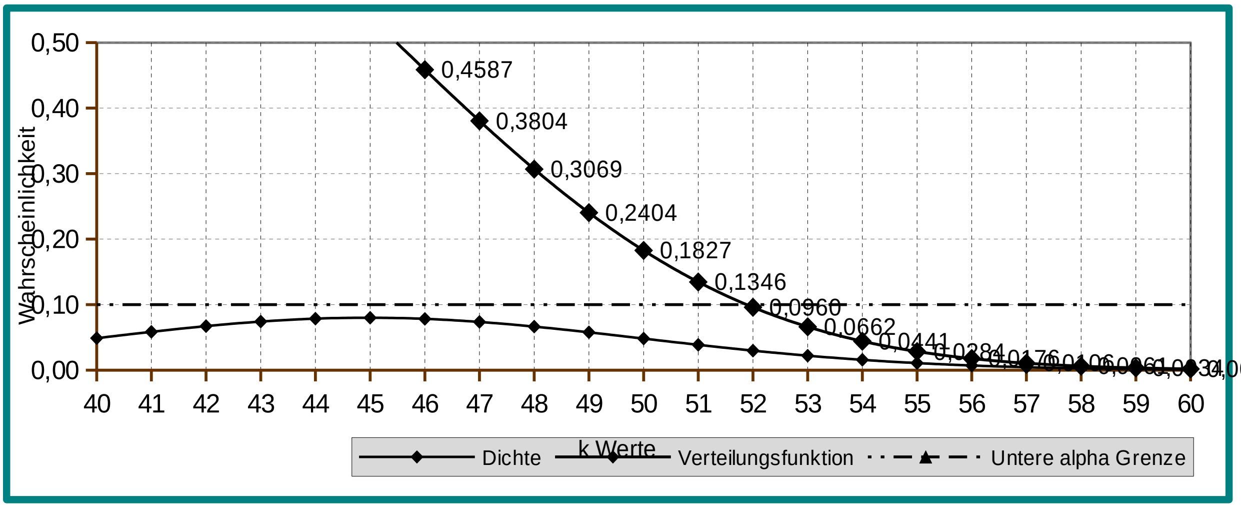
Warum steht beim rechtsseitigen Hypothesentest: $H_0: \; p \leq p_0$

Der Anbieter Elektronikteilen sagt: Sein Produkt hat eine Ausfallquote von 4% (p_0 = 4%). Wir als Nutzer haben Zweifel an der Quote und meinen er untertreibt : p > 4%. **Das ist die Hypothese H_1.** Wenn die Ausfallquote kleiner ist, haben wir Pech gehabt, also gehören alle Wahrscheinlichkeiten $p < p_0$ zur positiven Seite des Anbieters, damit gehören zur Anbieterseite: $p \leq p_0$. **Das ist die Hypothese H_0**, die wir widerlegen wollen, weil unsere Auffassung lautet: p > 4%. Das bedeutet, wenn bei einer Auswahl von Teilen sehr viele Teile ausfallen, dann spricht das gegen die Aussage des Anbieters. Damit liegt der Ablehnungsbereich rechts, bei großen Werten, damit ein Rechtsseitiger Hypothesentest.

H_0: **$p \leq 0{,}4$** und n = 100 und als Signifikanzniveau α = 0,1



Die Grenze des Ablehnungsbereichs für p = 0,4 liegt bei k = 47. Der Wert 47 selbst tritt nur mit einer Wahr-scheinlichkeit von 2,9% auf und alle größeren Werte noch seltener. Insgesamt treten alle Werte, die größer sind als 47 nur mit einer Wahrscheinlichkeit von 0,093 auf.



Bei einer Wahrscheinlichkeit von p = 0,45 treten Werte, die größer als 47 sind, mit einer Wahrscheinlichkeit von 38% auf, damit viel häufiger. Der Wert 47 selbst mit einer Wahrscheinlichkeit von 7,3 %. damit sprechen größere Werte als 47 wohl eher dafür, dass die tatsächliche Wahrscheinlichkeit größer als 4 % ist.

Mit der **größten angegebenen Wahrscheinlichkeit** ist der

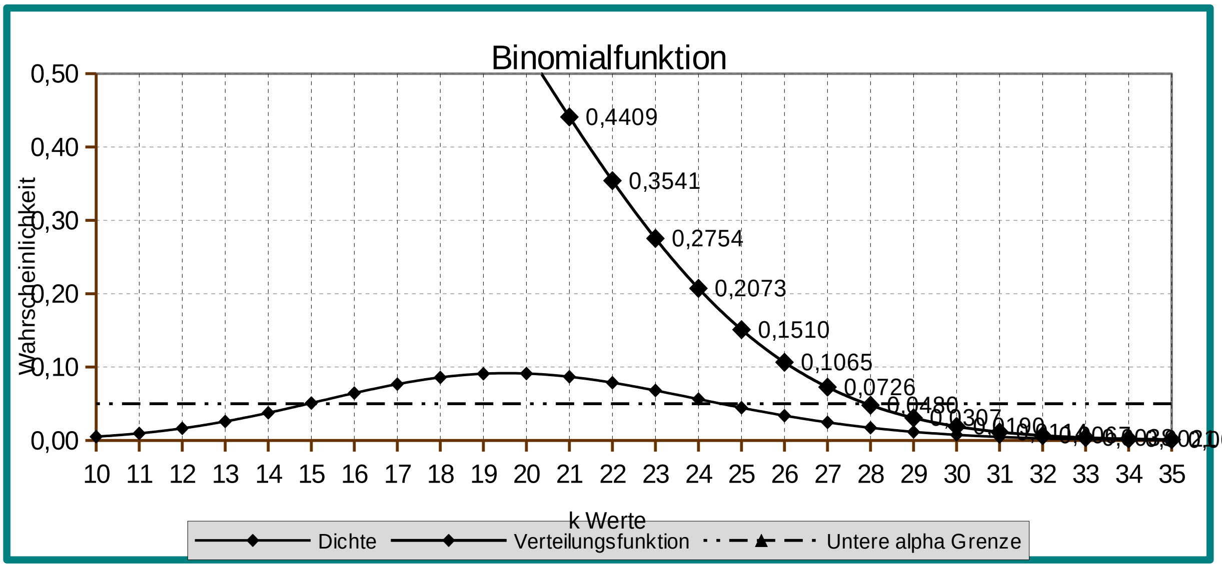
- Ablehnungsbereich so weit nach rechts verschoben, dass man bei noch größeren Werten eher Zweifel an der Wahrscheinlichkeit hat, als dass man eine ungünstige Stichprobe erhalten hat.
- der Annahmebereich bereits so weit nach rechts verschoben ist, dass man bei noch größeren Werten an der Wahrscheinlichkeit zweifelt.

Im angegebenen Fall liegt der Erwartungswert bei E(X) = n • p = 100 • 0,04 = 40. Man räumt die Möglichkeit ein, dass man zufällig eine ungünstige Probe zieht und lässt noch zu, dass mehr Teile als diese 40 ausfallen können. Aber eine noch größere Anzahl als 47 lässt dann doch Zweifel aufkommen.

Beispiel.
Ein Großhändler verkauft Glühbirnen an einen Baumarkt. Er gibt an, dass von den Glühbirnen (höchstens) **5%** defekt sein können. Der Baumarkt zieht eine Probe von **400 Glühbirnen** und will auf Grund dieser Stichprobe entscheiden, ob er eine größere Menge Glühbirnen bestellt. Wie könnte eine Entscheidungs-regel lauten, so dass der Baumarkt in **95%** der Fälle richtig liegt?

Das Signifikanzniveau
Das Signifikanzniveau ist 5%. Da der Baumarkt mit einer Defektrate unter 7% zufrieden ist, wird er beispiels-weise ein Ergebnis von 2% defekten Glühbirnen nicht als Ablehnungsgrund für die Lieferung ansehen. Der Ablehnungsbereich kann damit komplett auf einer Seite des Erwartungswertes liegen.

Die Art des Tests
Da man durch einen Test eine Hypothese nie beweisen kann, sondern sie bei einer großen Abweichung höchstens verwerfen kann, wird der Baumarkt für den Test davon ausgehen, dass mehr als 7% der Glühbirnen defekt sind. Man setzt die Angabe des Anbieters, dass muss nicht immer ein Produkt sein, mit einem Gleich-heitszeichen an. In diesem Fall p_0 = 7%. Als Abnehmer stellen wir uns auf den Standpunkt, die Angabe des Anbieters ist falsch (Er „lügt"). Dann ist für uns als Abnehmer der Fall uninteressant, wenn weniger als 7% defekte Glühbirnen vorhanden sind, für uns ist nur interessant, wenn mehr als 7% defekte Glühbirnen vorhan-den sind. Damit ist unsere Behauptung p > 7%. **Das ist die Hypothese H_1 !**
Damit bleibt für die Aussage des Anbieters als Erweiterung p ≤ 7%. **Das ist die Hypothese H_0**, der wir nicht trauen und die wir widerlegen wollen.
Kleinere Wahrscheinlichkeiten bringen eine Binomialverteilung hervor sich die nach links verschieben. Damit verschiebt sich der Erwartungswert (höchste Säule) und seine Umgebung (die Werte, die am häufigsten auftreten) auch nach links. Das stört uns nicht.
Größere Wahrscheinlichkeiten verschieben die Funktion nach rechts. Das heißt, größere Wahrscheinlich-keiten verschieben den Erwartungswert nach rechts, größere werte treten öfter auf. das wollen wir nicht, das ist für uns ungünstig. Also sagen wir: Große Werte k, die größer, als eine von uns festgelegtes k_0 sind, dürfen nur selten auftreten. Damit sollen Werte, die größer als k_0 sind insgesamt (alle noch größeren) höchstens mit einer Wahrscheinlichkeit von α (Signifikanzniveau) auftreten. Damit entsteht ein rechtsseitiger Test, da der Ablehnungsbereich am rechten Ende der Wahrscheinlichkeitsfunktion auftritt.



Das sind die zugehörigen Funktionskurven: Dichtefunktion (Einzelwahrscheinlichkeiten) und umgekehrte Verteilungsfunktion (1 – $F_{400;0,05}$(k-1)).
Interpretation der Kurven:
Bei der angegebenen Wahrscheinlichkeit von 7% sollten bei einer „normalen" Stichprobe etwa 19 oder 20 defekte Glühbirnen auftreten. Der Erwartungswert liegt bei E(X) = n • p = 400 • 0,05 = 20. Da man bei den Stichproben auch Pech haben kann und zufällig mehr defekte Glühbirnen findet, sind also auch Werte größer als 20 durchaus noch möglich. Betrachtet man die Einzelwahrscheinlichkeit k = 23, so kann diese immer noch mit einer Wahrscheinlichkeit von 6,8 % auftreten. Aber der Wert k = 28 nur noch mit einer Wahrscheinlichkeit von 1,7 %, alle größeren Werte **noch seltener** , dh.: möglich, aber eher unwahr-scheinlich. Deshalb legen wir fest, die Summe der unwahrscheinlichen Ergebnisse sollen insgesamt nicht mehr als 5 % betragen.
Größere Werte bedeuten, dass die Dichtefunktion weiter nach rechts rückt, das bedeutet aber, dass die Wahrscheinlichkeit größer als 7% beträgt. Ab k = 28 ist mit 0,048 der kritische Bereich erreicht.