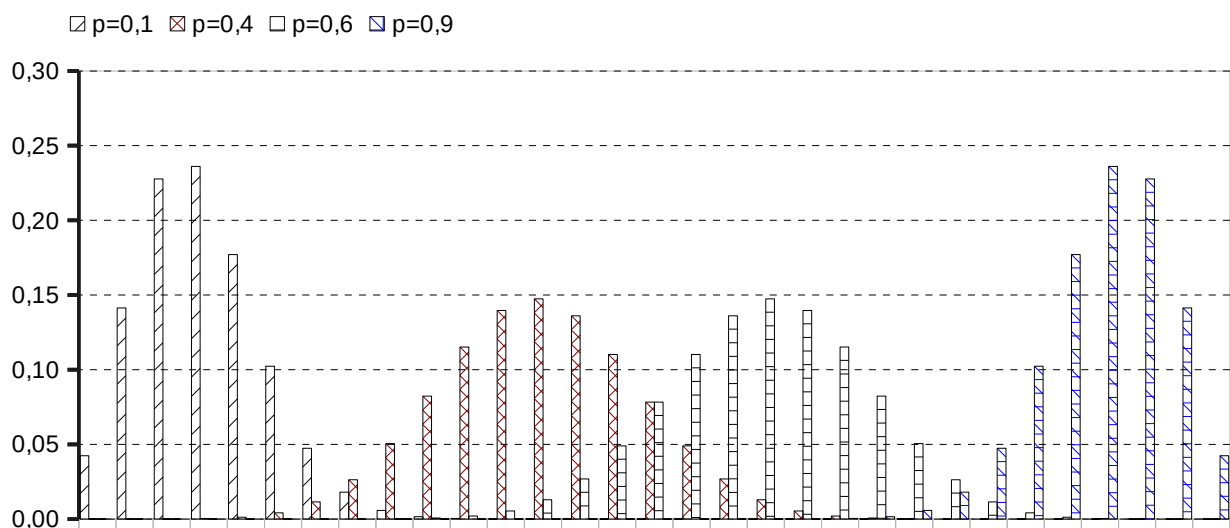


### 34. Hypothesentest

Die bisherigen Wahrscheinlichkeitsaufgaben waren so angelegt, dass man aus einem von vornherein bekannten Aufbau die Wahrscheinlichkeiten für ein bestimmtes Ereignis berechnet hat. Es war bekannt, welche Farbzusammensetzung die Kugeln in einer Urne haben und es wurde die Frage nach den Kugeln einer bestimmten Farbe gefragt.

Beim Hypothesentest steht die Frage andersherum: Man kennt ein Ziehungsergebnis und möchte von dem aus auf den Inhalt der Urne schließen. Das das natürlich sehr vage ist, ist klar. Bekannt sind noch grobe Verhältnisse der Farbanordnung in der Urne. Etwa in der Art, es gibt mehr schwarze als weiße, oder mindestens 60% der Kugeln sind weiße. Gleichgültig, welche Aussage man trifft, jede Aussage wird eine bestimmte Fehlerquote enthalten und man kann durch ein extrem unwahrscheinliches Ziehungsergebnis auf eine ganz falsche Fährte gesetzt werden.

Deshalb ist eine der zentralen Fragen in diesem Kapitel, mit welcher Sicherheit kann man die gewählte Aussage treffen.



In dem Bild sind verschiedene Binomialverteilungen dargestellt, bei denen immer der Wert für  $n = 30$  ist. Es wechselt nur die Wahrscheinlichkeit. Die einzelnen Wahrscheinlichkeiten sind am oberen Rand dargestellt. Daraus ist zu erkennen, dass mit wachsender Wahrscheinlichkeit die Säulen nach rechts wandern. Es sollen jetzt die roten Säulen im Zusammenhang mit dem Hypothesentest betrachtet werden. Die Wahrscheinlichkeit für die roten Säulen ist 0,4.

Es wird angenommen, jemand liefert irgendwelche Teile aus und sagt :

**Mindestens 40%** meiner Teile sind fehlerfrei.

Der Empfänger der Teile macht eine Stichprobe und entnimmt der Lieferung 30 Teile. Damit hat man genau die obige Situation mit  $p = 0,4$ . Nach den Regeln der Wahrscheinlichkeit ist dann zu erwarten, daß **mindestens 12 Teile** fehlerfrei sind, da 12 der Erwartungswert von  $n=30$  und  $p = 0,4$  ist. Sind es mehr fehlerfreie Teile, kann der Empfänger nur zufrieden sein, aber wie sieht es umgekehrt aus. Was ist, wenn man nur 9 fehlerfreie Teile findet. Aus dem Säulendiagramm kann man entnehmen, daß 9 Teile bei einer Wahrscheinlichkeit von 40% gar nicht so selten sind. Möglicherweise hat man eine ungünstige Stichprobe gezogen und die 40% sind doch richtig. Was ist aber, wenn nur 5 Teile fehlerfrei sind. Bei einer Wahrscheinlichkeit von 40 % sind 5 fehlerfreie Teile

eher selten. Man kann wohl eher davon ausgehen, daß die 40% fehlerfreie Teile nicht der Tatsache entsprechen und die Wahrscheinlichkeit für fehlerfreie Teile eher geringer ist, für kleinere  $p$  ist die Säule an  $k = 9$  größer. Das bezeichnet man dann als **Linksseitigen Test**, weil man die linke Grenze testet, ob diese eingehalten wird. Die rechte Grenze interessiert den Empfänger nicht, denn wenn die Anzahl fehlerfreier Teile größer ist, kann das für ihn nur positiv sein.

Wie sieht es aus, wenn man die Sache umgekehrt betrachtet. Der Lieferant behauptet **Höchstens 40%** meiner Teile sind fehlerhaft.

(Hier geht es nicht um Produktionsprozesse, sondern um Wahrscheinlichkeit)  
Der Empfänger der Teile macht eine Stichprobe und entnimmt der Lieferung 30 Teile. Wieder muß er damit rechnen, daß er in der Stichprobe 12 defekte Teile findet. Findet er in seiner Stichprobe nur 6 oder 9 defekte Teile, ist das für ihn positiv und interessiert ihn nicht. Die Säulen der Binomialverteilung sagen ihm aber auch, daß er durchaus auch mit 15 oder 16 defekten Teilen rechnen kann, wenn er eine ungünstige Stichprobe gezogen hat. Die Wahrscheinlichkeit für so viele fehlerhafte Teile ist durchaus noch zu beachten. Findet er aber 19 oder 20 Teile sieht das anders aus. Bei einer Wahrscheinlichkeit von 40% sind so viele Teile doch eher selten und es wäre ein großer Zufall, wenn er gerade eine solche ungünstige Probe gezogen hätte, daß 19 defekte Teile gerechtfertigt sind. In diesem Fall muß man wohl eher davon ausgehen, daß der Prozentsatz fehlerhafter Teile größer ist. Hier spricht man von einem **Rechtsseitigen Test**, da den Tester nur die rechte Grenze interessiert, aber nicht die linke.

In beiden Fällen muß man aber konstatieren: Die extrem abweichenden Zahlen bei den Tests sind trotzdem nicht ausgeschlossen. Auch für diese Abweichungen gibt es noch eine Wahrscheinlichkeit. Man bemängelt die Lieferung, obwohl die Lieferung den Aussagen des Produzenten entspricht. Man hat tatsächlich eine solche schlechte Probe erwischt, oder die Teile der Maschine, die hauptsächlich schlechte Teile produziert, lagen gerade oben drauf. Man kann sich also auch bei seiner Einschätzung irren. Das ist die Irrtumswahrscheinlichkeit mit der man leben muß.

Grundsätzlich unterscheidet man beim Hypothesentest,

- ob es sich um die **Bestätigung oder Verwerfung einer Hypothese** handelt (einseitiger Hypothesentest) oder
- ob man **zwei Hypothesen gegenüberstellt**, um diejenige zu finden, die wahrscheinlicher ist (zweiseitiger oder Alternativer Hypothesentest)

Gibt es nur eine Hypothese unterscheidet man zwischen einem linksseitigen und einem rechtsseitigem Hypothesentest, je nachdem, wo der **Ablehnungsbereich** liegt.

Es kommt darauf an, Kriterien zu finden, nach denen die Hypothese abzulehnen ist oder nicht abgelehnt werden kann.

**Eine Hypothese wird grundsätzlich nicht angenommen.**

Wenn eine Hypothese nicht abgelehnt werden kann, heißt das noch lange nicht, daß sie richtig ist. Mit diesem Test kann man die Hypothese nicht ablehnen. Das bedeutet ja nicht, daß man vielleicht mit einem anderen oder besseren Test mehr Einzelheiten erfahren kann und sich dann herausstellt, daß die Hypothese doch falsch ist. Wenn aber die Hypothese schon diesen Test nicht übersteht, dann ist wohl doch eher davon auszugehen, daß sie falsch ist und kann abgelehnt werden. Dabei bleibt die Irrtumswahrscheinlichkeit immer noch offen.

Es ist aber auch möglich, einen zweiseitigen Test durchzuführen, ohne das eine zweite Hypothese im Spiel ist. Damit möchte man erreichen, daß die Ergebnisse weder zu

weite niedriger sind als der Erwartungswert, noch daß sie zu weit nach oben abweichen.

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

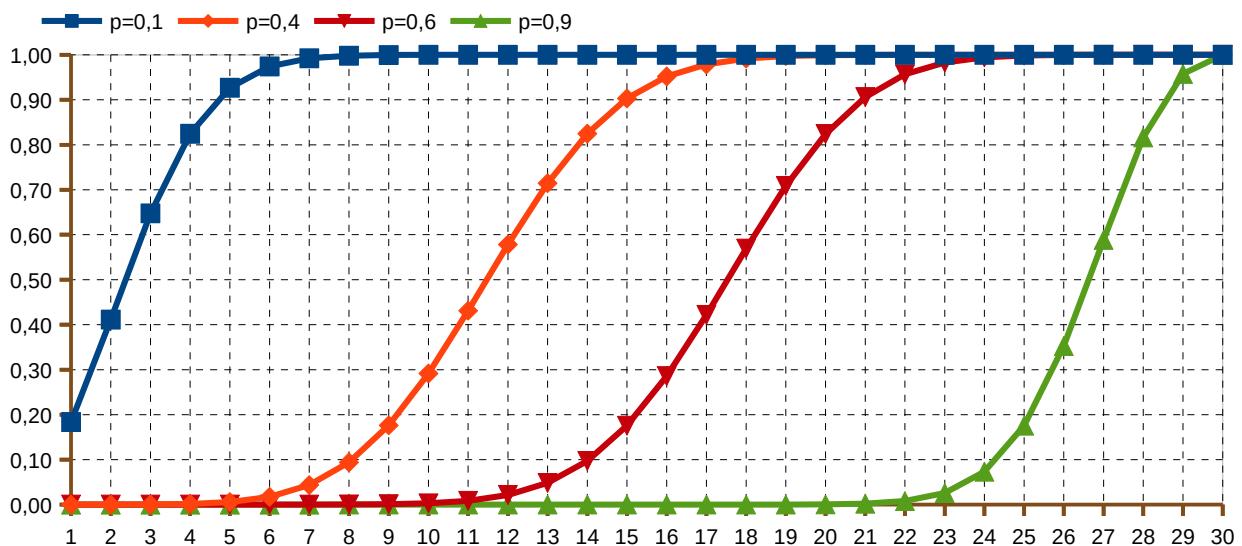
Bei einem rechtsseitigen Test wird die Hypothese aufgestellt

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*)

Beim Hypothesentest wird aber nicht mit den Einzelwahrscheinlichkeiten gearbeitet, sondern mit den summierten Wahrscheinlichkeiten. Die dabei auftretende Irrtumswahrscheinlichkeit ist die Summe aller Wahrscheinlichkeiten, die den Ablehnungsbereich ausmachen. Für die obigen Wahrscheinlichkeiten von 0,1 ; 0,4 ; 0,6 und 0,9 sehen die summierten Wahrscheinlichkeiten für  $n = 30$  folgendermaßen aus:



Auch hier zeigt sich die Erscheinung, je größer die Wahrscheinlichkeit, desto weiter rechts, zum  $n$  hin, liegen die Kurven.

Es soll zunächst ein linksseitiger Test betrachtet werden, dh. der Ablehnungsbereich liegt am linken Ende der Funktion und als Irrtumswahrscheinlichkeit werden  $10\% = 0.1$  angenommen. Für die Wahrscheinlichkeit von 0,4 (die Kurve, mit den auf die Spitze gestellten Quadraten) wird bei  $k = 8$  der summierte Wert 0.1 erreicht. In diesem Fall endet der Ablehnungsbereich bei  $k = 7$ .

Betrachtet man einen rechtsseitigen Test für die Wahrscheinlichkeit 0,4 und eine Irrtumswahrscheinlichkeit von 0.1, dann sind die Funktionen an der oberen Grenze zu untersuchen. Eine rechtsseitige Irrtumswahrscheinlichkeit von 0.1 bedeutet, dass zum Annahmehbereich alles gehört bis zu einer summierten Wahrscheinlichkeit von 0.9. Von 0.9 bis 1 beliebt an der oberen Grenze der summierten Wahrscheinlichkeit ein Bereich von 0.1 übrig, der den rechtsseitigen Ablehnungsbereich angibt. Bei der obigen Kurve von  $p = 0,4$  ist das bei  $k = 15$  der Fall.

Bei einem rechtsseitigen Hypothesentest ist das erste  $k$ , das die Grenze von 0.9 überschreitet das Ende des Annahmebereichs und nicht der Beginn des Ablehnungsbereichs. Zur Begründung siehe Abschnitt „Rechtsseitiger Hypothesentest“.

Auch Ereignisanzahlen, die über/unter der vorgegebenen 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  **$\alpha$  - Fehler**

Die Wahrscheinlichkeit für den Fehler ist genau die Summe der Wahrscheinlichkeiten, mit denen das Ereignis eintreten kann, bei dem die Anzahl  $k$  größer oder kleiner 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 !

Woran erkennt man, was die  $H_0$ -Hypothese und was die  $H_1$  Hypothese ist ?

Merke beim Signifikanztest (gilt nicht für Alternativtest):

In der  $H_1$  - Hypothese steht niemals ein  $=$ ,  $\leq$  oder  $\geq$

### Das Gleichheitszeichen gehört immer zu $H_0$ !

Wenn in der Aufgabe das Wort **höchstens** ( $p \leq p_0$ ) auftaucht, dann wissen wir, dass das die  $H_0$ -Hypothese kennzeichnet und wir einen **rechtsseitigen** Hypothesentest durchführen.

Wenn in der Aufgabe das Wort **mindestens** ( $p \geq p_0$ ) auftaucht, dann wissen wir, dass das die  $H_0$ -Hypothese kennzeichnet und wir einen **linksseitigen** Hypothesentest durchführen.

Wenn in der Aufgabe das Wort **mehr als** oder auch **größer** ( $p > p_0$ ) auftaucht, dann wissen wir, dass das die  $H_1$  – Hypothese kennzeichnet und wir einen **rechtsseitigen** Hypothesentest durchführen.

Wenn in der Aufgabe das Wort **weniger als** oder auch **weniger** ( $p < p_0$ ) auftaucht, dann wissen wir, dass das die  $H_1$  – Hypothese kennzeichnet und wir einen **linksseitigen** Hypothesentest durchführen.

Generell stellt man als erstes eine sogenannte Nullhypothese auf. Das ist etwa die These, die man am wahrscheinlichsten hält. Diese wird bezeichnet mit  $H_0$ . Arbeitet man mit einer Gegenhypothese, wird diese als  $H_1$  bezeichnet. Bei statistischen Tests geht es darum, mit einem Entscheidungsverfahren die Ergebnisse einer Stichprobe auszuwerten. Die zu beantwortende Frage lautet:

„Muss aufgrund der Stichprobenergebnisse eine Hypothese (Nullhypothese  $H_0$ ) über die unbekannte Grundgesamtheit abgelehnt werden?“

Dass diese Nullhypothese so gewählt werden soll, dass sie abgelehnt werden soll hat mit den beiden möglichen Irrtümern zu tun und soll hier erst mal nicht im Mittelpunkt stehen. Bei der Beantwortung dieser Frage geht man immer ein gewisses Unsicher-

heitsrisiko ein. Durch Wahl entsprechender Annahme- oder Ablehnungsbedingungen kann man aber diesen Fehler in Grenzen halten.

Wenn in der Aufgabenstellung nicht konkret angegeben ist, mit welcher Nullhypothese man arbeiten soll, muß man die Hypothese, die bestätigt werden soll in  $H_1$  unterbringen und das Gegenteil in  $H_0$ . Das ist etwas gewöhnungsbedürftig und hat was mit doppelter Verneinung zu tun. Da man eine Hypothese  $H_0$  nicht bestätigen kann, wird die zu bestätigende Hypothese  $H_1$ . Wenn man nämlich  $H_0$  ablehnt, hat man  $H_1$  bestätigt, aber nicht umgekehrt. Also formuliert man die Hypothesen so, daß  $H_0$  abgelehnt wird, weil dann  $H_1$  bestätigt ist.

### 34.1 Der Alternativtest

Fällt die Entscheidung zwischen zwei Hypothesen, nennt man den Test Alternativtest. (Null)Hypothese  $H_0: p=p_0$  Alternative (Gegenhypothese)  $H_1: p=p_1$

Man hat also hier 2 feste Verteilungen, die miteinander verglichen werden.

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  $x_1, x_2, \dots, x_n$  besteht nun aus der Angabe eines kritischen Bereichs  $K$  (Verwerfungsbereich) zusammen mit der Entscheidungsregel:

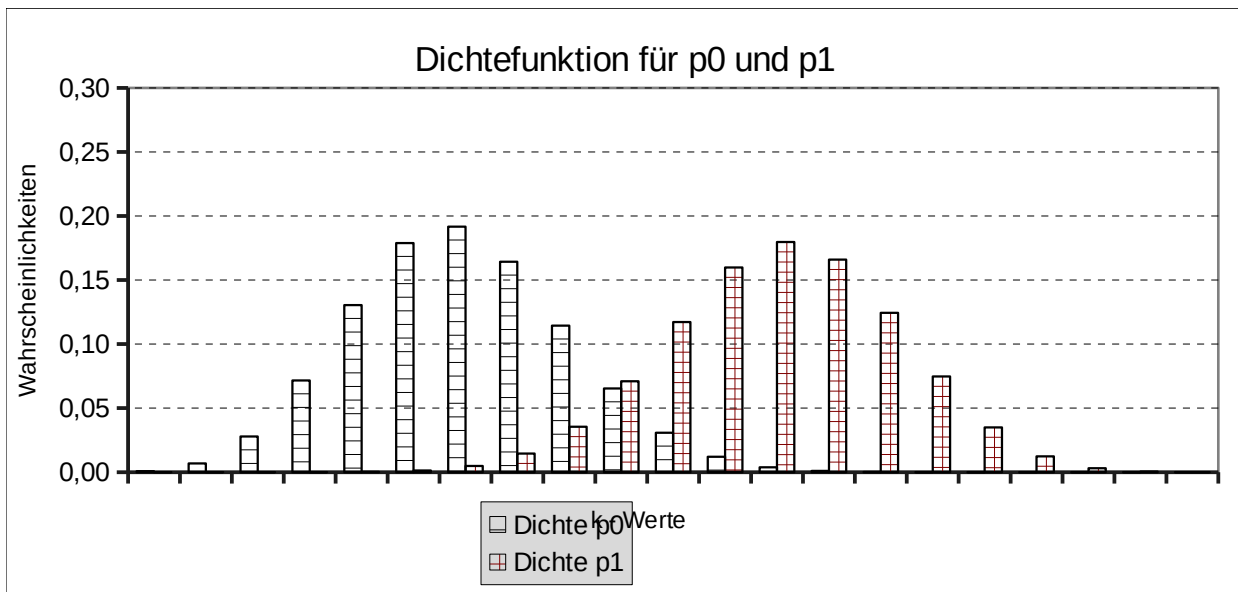
- Verwerfe die Nullhypothese, falls die Werte im kritischen Bereich liegen.
- Belasse die Nullhypothese, falls nicht

Ein Belassen der Nullhypothese bedeutet nicht, daß die Nullhypothese richtig ist, sondern nur, daß sie nicht widerlegt werden kann !

Der Grundsatz bei Alternativtests lautet in etwa folgendermaßen: Es gibt eine länger zurückliegende Untersuchung, die ein bestimmtes Ergebnis gebracht hat. z.B. Die Anzahl fehlerhafter Teile, die Wirksamkeit eines Medikament u.ä. Das ist die alte Hypothese (=Aussage). Seit dieser Zeit sind Veränderungen durchgeführt worden. Es soll nun geprüft werden, ob diese Veränderungen eine Änderung er vorherigen Situation gebracht hat. Das ist die neue Hypothese, in Testverfahren als Nullhypothese beschrieben.

In beiden Hypothesen geht es um die Wahrscheinlichkeit, mit der das Ereignis eintritt. Während man bei der alten Hypothese mehr oder weniger gesicherte Werte hat, da diese vielleicht über Jahre sich bestätigt haben, kennt man die Wahrscheinlichkeit der neuen Aussage nicht. Man spricht für diese Wahrscheinlichkeit eine Vermutung aus. Jetzt kann man durch Probeentnahmen feststellen, ob sich der Wert bestätigen lässt, oder nicht. Dabei gibt es ein grundlegendes Problem: Die Probenentnahme bleibt eine eingeschränkte zufällige Sicht auf das Ereignis. Sie liefert weder die Bestätigung der Annahme noch die Verneinung der Annahme. Aber durch eine mehrfache Probenentnahme kann man dem tatsächlichen Ergebnis schon sehr nahe kommen.

Daraus resultiert die Aufgabestellung: Mit welcher Sicherheit kann man bei einer Probenentnahme von  $n$  Objekten die vermutete Veränderung bestätigen oder als falsch verwerfen.



Man geht bei solchen Untersuchungen von einer Binomialverteilung aus, auch wenn die Hypergeometrische Verteilung theoretisch angebrachter wäre. Dabei legt man eine sehr große Anzahl der Gesamtmenge zu Grunde, so dass das Ziehen ohne Zurücklegen wie ein Ziehen mit Zurücklegen wirkt. Außer dem Zentralen Grenzwertsatz, siehe weiter vor bei Normalverteilung, gilt auch eine andere Beziehung: Mit steigendem  $n$  nähert sich die Hypergeometrische Verteilung der Binomialverteilung an.

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  $k=9$  (oder  $10$ ). Die Alternativhypothese von  $k=3$  bis  $k=17$ .

Was bedeutet das für den Test:

Werte von  $k > 11$  werden wohl nicht auftreten, denn dann wäre die Summe der Wahrscheinlichkeiten von  $H_0$  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 Werten von  $k=7$  bis  $k=10$ . Ist diese Anzahl  $H_0$  oder  $H_1$  zuzuordnen.

Und an dieser Stelle kommt genau der  $\alpha$ -Fehler zum Tragen:

Legt man die Grenze für  $k_0$  zur Bestätigung oder Verwerfung von  $H_0$  auf den Wert  $k=10$ , dann gehören die Werte  $k > 10$  zum  $\alpha$ -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  **$\alpha$ -Fehler** oder **Fehler 1. Art**.

Was ist andererseits mit  $k=8$ .

Bei  $k=8$  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  $k=7$  und  $k=6$  möglich. In diesem Fall spricht man von einem  $\beta$ -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  **$\beta$  - Fehler** oder **Fehler 2. Art**.

$\beta$  - Fehler setzen immer eine zweite Hypothese voraus und sind für das Abitur nicht relevant.

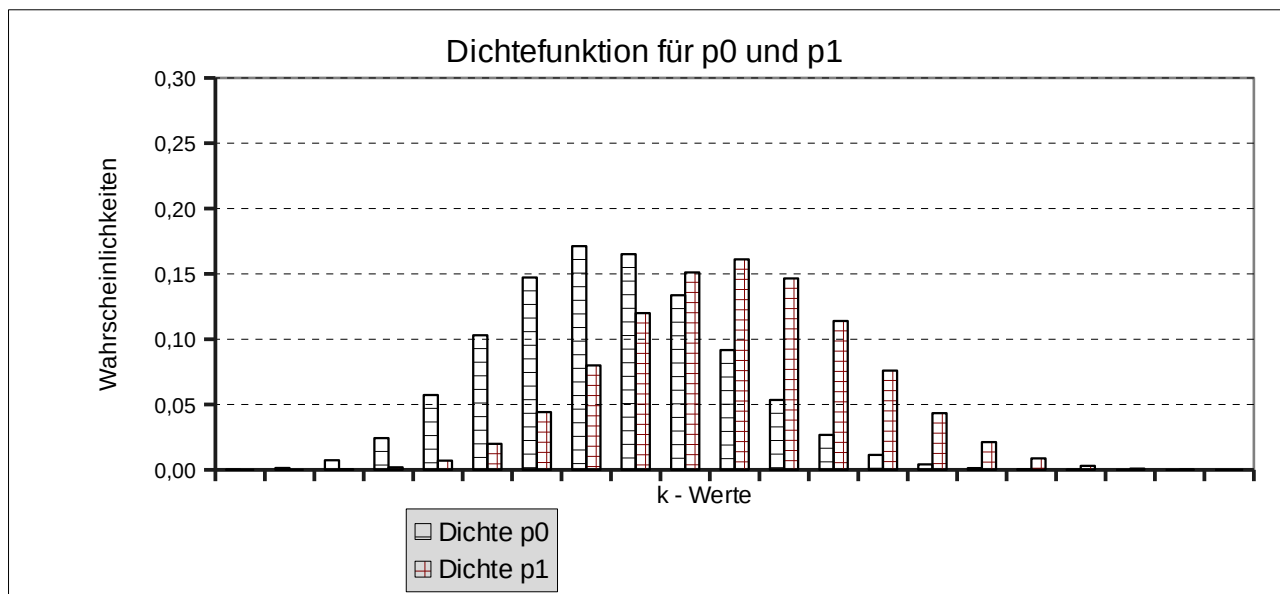
Die Nullhypothese und Alternative sind nicht gleichwertig:

- 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:

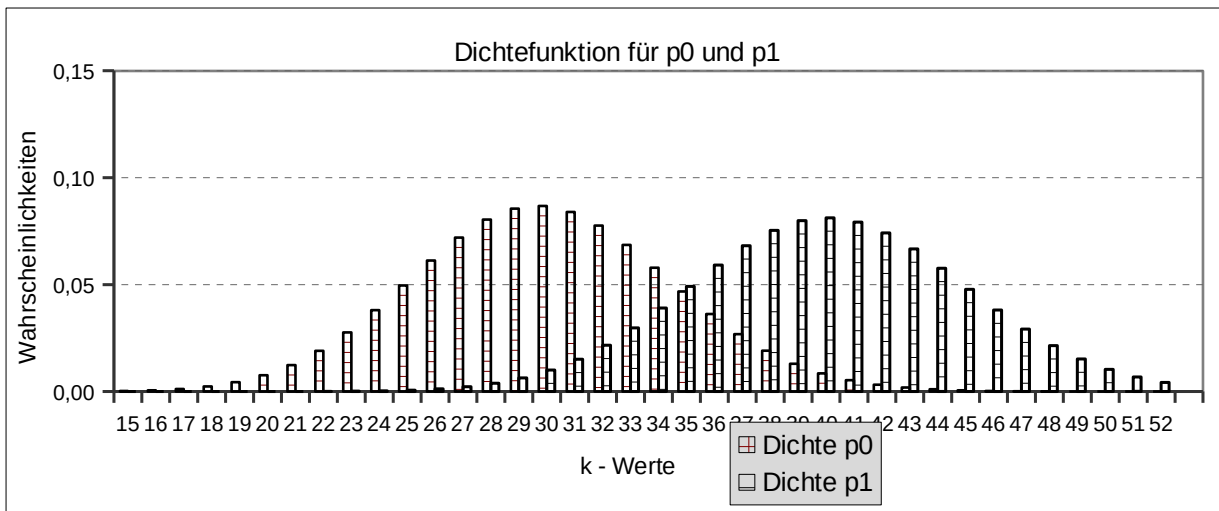
$$P_{H_0} [(x_1, x_2, \dots, x_n) \text{ liegt im Verwerfungsbereich}] \leq \alpha.$$

Das bedeutet für einen rechtsseitigen Test, daß die aufsummierte Wahrscheinlichkeit ab einem Wert  $k_0$  kleiner als  $\alpha$  sein muss und für einen linksseitigen Test, daß die aufsummierte Wahrscheinlichkeit bis zu einem Wert  $k_0$  kleiner als  $\alpha$  sein muß.



Aus den beiden Dichtefunktionen 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 Wahrscheinlichkeiten 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  $H_0$  und  $H_1$
2. Je kleiner der Stichprobenumfang, desto schlechter die Trennung in  $H_0$  und  $H_1$

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  $\alpha$ -Fehler verkleinern will, erhöht man den  $\beta$ -Fehler und umgekehrt.

Je größer das  $k$ , desto kleiner der  $\alpha$ -Fehler. und desto größer der  $\beta$ -Fehler

Je kleiner das  $k$ , desto kleiner der  $\beta$ -Fehler und desto größer der  $\alpha$ -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.

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

$H_0$  : "Die Schachtel ist hoher Qualität"

Nullhypothese

$H_1$  : " 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"

$0 \leq X \leq 3$  : Die Schachtel ist gute Qualität

Entscheidung für  $H_0$  gegen  $H_1 \Rightarrow$  Annahmehbereich für  $H_0$

$3 \leq X \leq 10$  : Die Schachtel ist minderer Qualität

Entscheidung für  $H_1$  gegen  $H_0 \Rightarrow$  Annahmehbereich von  $H_1$  bzw.

Ablehnungs-/ kritischer Bereich von  $H_0$

**(3) Mögliche Fälle beim Prüfen**

Realität	Entscheidung aufgrund des Testverfahrens	
	$H_0$ ist wahr ( $p_0 = 0.1$ )	$0 \leq X \leq 3$ Annahme von $H_0$ richtige Entscheidung
$H_1$ ist wahr ( $p_1 = 0.4$ )	$0 \leq X \leq 3$ Ablehnung von $H_0$ falsche Entscheidung <b>Fehler 2. Art</b>	$3 \leq X \leq 10$ 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:  $n = 10; \quad p_0 = 0,1; \quad k > 3$   
 $P(X > 3) = 1 - F_{10;0,1}(3) = 0,01279$

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:  $n = 10; \quad p_1 = 0,4; \quad k < 3$   
 $P(X \leq 3) = F_{10;0,1}(3) = 0.38228$

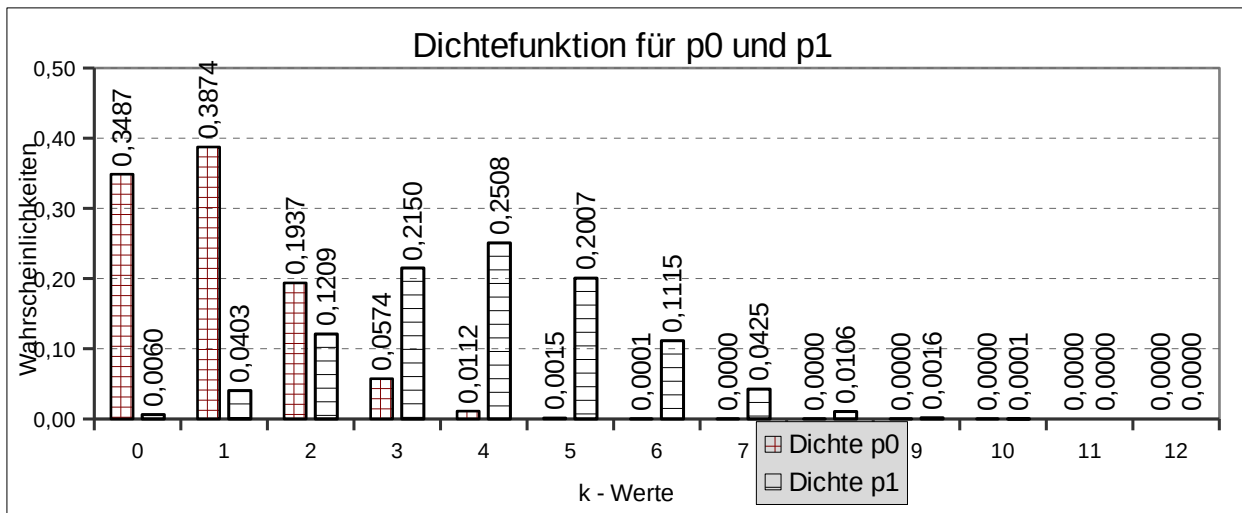
Die Wahrscheinlichkeit, dass die Kiste 2. Wahl enthält (d.h.  $H_1$  richtig ist), aber sie trotzdem ablehnt ist 38,23%.

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



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  $\alpha$ -Fehler die Wahrscheinlichkeit ab der Position  $k = 4$  aufzuaddieren.

k	P(X=k)
4	0.01116
5	0.00149
6	0.00014

Summe: 0,01279

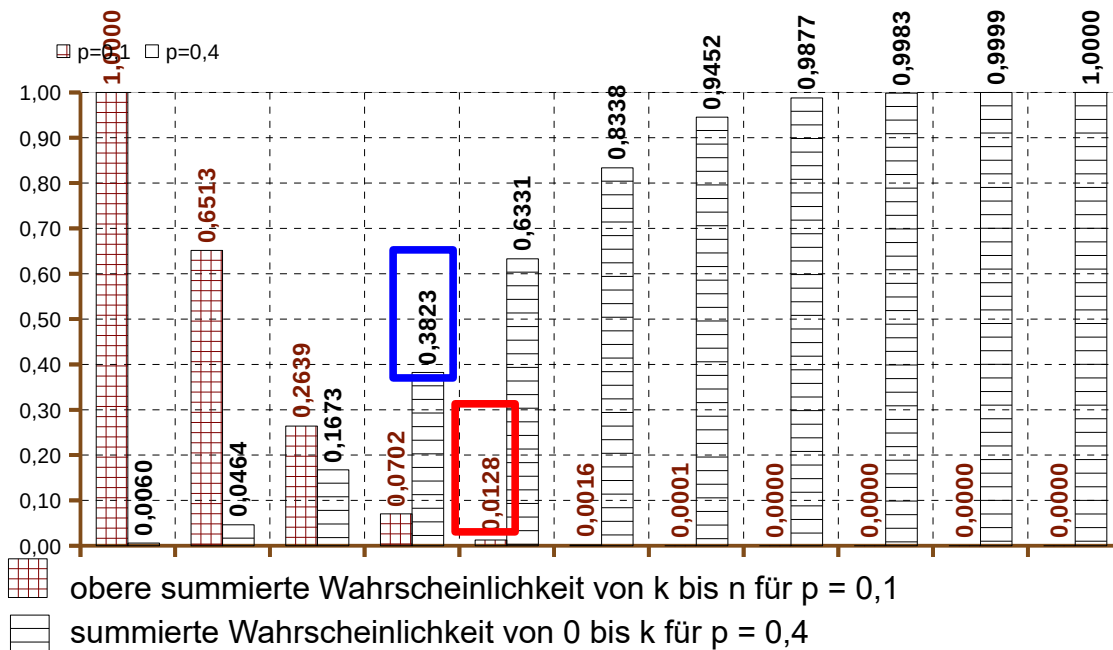
(s. Wert für den  $\alpha$ -Fehler auf der vorhergehenden Seite)

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  $\beta$ -Fehler die Wahrscheinlichkeit bis zur Position  $k = 3$  aufzuaddieren.

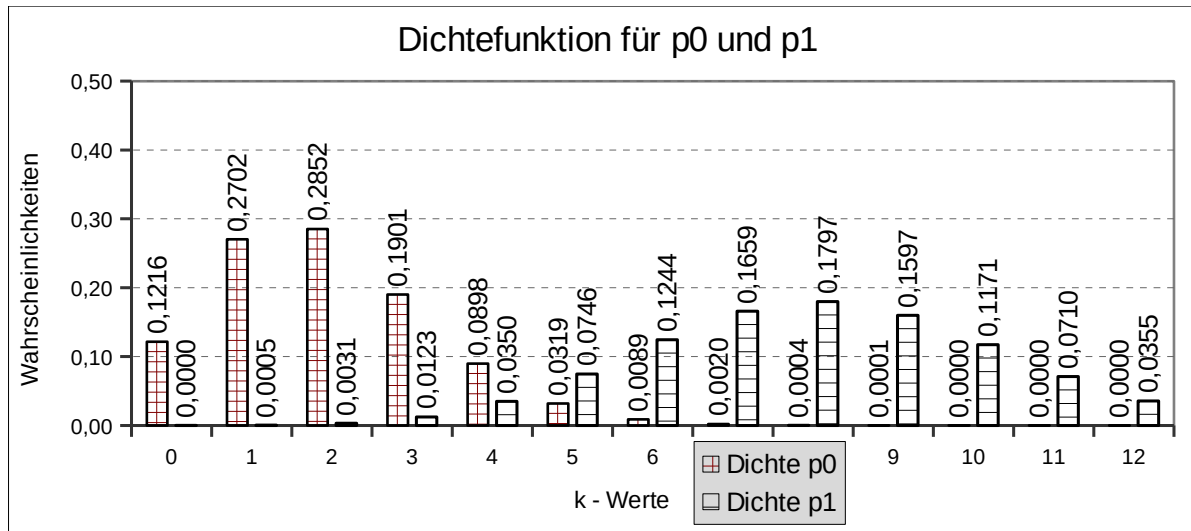
k	P(X=k)
0	0.0060
1	0.0403
2	0.1209
3	0.2150

Summe: 0,3822

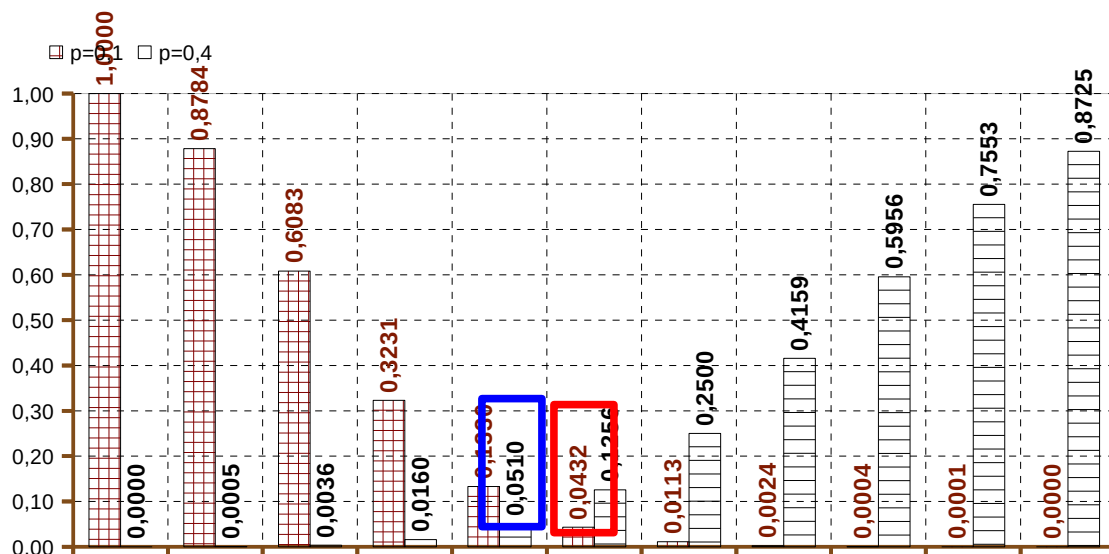
(s. Wert für den  $\beta$ -Fehler auf der vorhergehenden Seite)



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.



Der  $\alpha$  - Fehler liegt jetzt bei 0,04317 und der  $\beta$  - Fehler bei 0,05095. Auf die Berechnung soll hier nicht noch einmal eingegangen werden, aber die Werte an einer summierten Binomialverteilung dargestellt werden. Dabei geht es nur um den Ausschnitt, der für die Entscheidung relevant ist.



Bei 4 schlechten Schrauben wird die Schraubenkiste noch als 1. Qualität angenommen. Dabei begeht man einen Fehler von 5%:  $H_0$  wird angenommen, obwohl sie falsch ist  $\rightarrow \beta$  Fehler;

Ab 5 schlechten Schrauben stuft man die Kiste unter mindere Qualität ein.  $H_0$  wird abgelehnt, obwohl die Kiste gute Qualität hat  $\rightarrow \alpha$  Fehler. Bei  $k = 5$  und  $p = 0,1$  verbleibt eine summierte Wahrscheinlichkeit bis  $n$  von 0,0432.

## Zusammenfassung

	Entscheidung aufgrund der Stichprobe	
	Annahme von $H_0$ ; Ablehnung von $H_1$	Annahme von $H_1$ ; Ablehnung von $H_0$
$H_0$ trifft zu mit $p = 0,2$	Richtige Entscheidung, <b>Sicherheit des Urteils</b> $P(A) = F^{100}_{0,2}(X \leq 25) = 1 - \alpha = 0,913$	Falsche Entscheidung, <b>Risiko 1. Art</b> $P(\bar{A}) = F^{100}_{0,2}(X \geq 26) = \alpha = 0,087$
$H_1$ trifft zu mit $p=0,3$	Falsche Entscheidung, Risiko 2. Art $P(\bar{B}) = F^{100}_{0,3}(X \leq 25) = 1 - \beta = 0,163$	Richtige Entscheidung <b>Sicherheit des Urteils</b> $P(B) = F^{100}_{0,3}(X \geq 26) = \beta = 0,837$

### 34.2 Der Signifikanztest

In den meisten Fällen gibt es jedoch keine zwei Alternativen, zwischen denen auszuwählen ist. Meist gibt es eine Hypothese, deren Wahrscheinlichkeit durch eine Stichprobe verifiziert oder widerlegt werden soll. Ein solcher Test ist der **Signifikanztest**.

Beim Testen will man mit einer Stichprobe vom Umfang  $n$  eine Hypothese  $H_0$  (z.B.  $p_0=70\%$ ) widerlegen ! Man geht dabei aus von einer Binomialverteilung  $Bn;p_0$  ( $k=0;1; \dots ;n$ ).

Das Signifikanzniveau (Verwerfungsbereich  $V$  der Nullhypothese) wird entweder mit  $\alpha$  (meist  $\alpha \leq 5\%$ ) oder aber mit einer Schranke  $k$  ( $V=\{0; \dots ;k\}$  bzw.  $V= \{k; \dots ;n\}$ ) festgelegt.

$V$  ist ein Bereich (für die  $k$ -Werte) in der Binomialverteilung  $Bn;p_0$  (je nach Lage der Alternativwahrscheinlichkeit  $p_1$  liegt  $V$  entweder links oder rechts bezüglich  $H_0$ ).

Die obere Schranke  $\alpha$  für das **Risiko 1. Art**, das man möglichst klein halten will, nennt man **Signifikanzniveau** (auch Irrtumswahrscheinlichkeit). Man legt damit fest, welchen Fehler 1. Art (d.h. welche Irrtumswahrscheinlichkeit für die aufgestellte Nullhypothese) man höchstens in Kauf nehmen möchte. Signifikanzniveau und Ablehnungsbereich der Nullhypothese hängen demnach zusammen.

$H_0$  wird abgelehnt, wenn das Stichprobenergebnis in **signifikanter** (bedeutsamer) Weise  $H_0$  widerspricht.

Die Wahrscheinlichkeit dieser **Nullhypothese  $H_0$**  ist  $p_0$ . Diese Hypothese wird als wahr angenommen, wenn das Ergebnis einer Stichprobe innerhalb eines bestimmten **Annahmebereichs** liegt. Liegt das Ergebnis jedoch im **Verwerfungsbereich**, ist die Hypothese aufzugeben. Natürlich gibt es auch hier wiederum die Möglichkeit, dass aufgrund der stochastischen Eigenschaften einer Stichprobe **Fehler** in der Einschätzung auftauchen.

- So kann z.B. die Nullhypothese aufgrund der Stichprobe abgelehnt werden, obwohl sie eigentlich zutrifft, wenn das Ergebnis (zufälligerweise) im Verwerfungsbereich zu finden ist (**Fehler 1. Art**, analog zum Alternativtest).
- Ebenso ist es denkbar, dass die Hypothese angenommen wird, da das Ergebnis (zufälligerweise) im Annahmehereich gelandet ist, obwohl die Hypothese gar nicht zutrifft (**Fehler 2. Art**, analog zu oben).

Die sogenannte **Irrtumswahrscheinlichkeit  $\alpha$**  ( $= P(\text{Fehler 1. Art})$ ) ist natürlich von der Größe des Annahmehereichs abhängig.

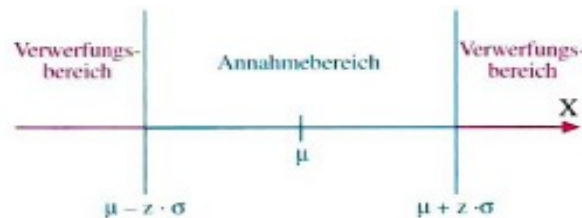
Als Nullhypothese sollte immer die Hypothese gewählt werden, deren irrtümliche Ablehnung als der schlimmere der beiden Fehler angesehen wird d.h. irrtümliche Verwerfung der (in Wirklichkeit wahren) Nullhypothese.

In vielen Fällen wird die Nullhypothese aber auch vorgegeben und es sollen die Angaben eines Produktherstellers aus dessen Sicht getestet werden.

Die untere Schranke  $1 - \alpha$  für die Wahrscheinlichkeit, eine wahre Hypothese zu Recht nicht abzulehnen, ist ein Maß für die Sicherheit der Entscheidung und heißt **Sicherheitsniveau  $1 - \alpha$** .

#### 34.2.1 Zweiseitiger Signifikanztests

Bei einem **zweiseitigen Signifikanztest** liegt der Annahmehereich (bei einer Binomialverteilung mit genügend großem  $n$ ) symmetrisch um den Erwartungswert  $\mu$ .



Da es sich hier um einen zweiseitigen Test handelt, wird der Annahmehereich so gelegt, dass rechts und links des Erwartungswertes die Hälfte der Irrtumswahrscheinlichkeit auftritt.

Ein (zweiseitiger) Signifikanztest läuft fast immer nach dem gleichen Schema ab:

1. Wie lautet die Nullhypothese ?
2. Wie groß sind Stichprobenumfang  $n$  und Signifikanzniveau  $\alpha$  ?
3. Bestimme den Annahmehereich (bzw. die Verwerfungsbereiche)
4. Entscheidung: Hypothese wird beibehalten, wenn Stichprobenergebnis innerhalb des Annahmehereichs liegt, ansonsten abgelehnt.

#### 34.2.2 Einseitige Signifikanztests

Annahmehereich  $A$  und Ablehnungsbereich  $\bar{A}$  der Hypothese  $H_0$  bestehen jeweils aus einem einzigen Intervall. Je nach Lage von  $A$  nennt man diese Tests rechts- bzw. linksseitig.

Zweiseitige Signifikanztests:  $\bar{A}$  wird durch  $A$  oder  $A$  wird durch  $\bar{A}$  in zwei Intervalle zerlegt.

Signifikanztests werden im Allgemeinen dann durchgeführt, wenn die Trefferwahrscheinlichkeit  $p$  unbekannt ist und deshalb eine über  $p$  gemachte Behauptung überprüft (getestet) werden soll.

### 34.2.3 Linksseitiger Hypothesentest mit Binomialverteilung

Bei einem linksseitigen Hypothesentest sprechen kleine Werte der Zufallsvariablen gegen die Hypothese, also Werte, die links auf dem Zahlenstrahl bzw. weit links vom Erwartungswert liegen.

Jemand behauptet, sein Ereignis (fehlerfreie Teile, Stimmenanteil ) hat eine Wahrscheinlichkeit von **mindestens** 40%:

$$H_0: p \geq p_0 = 40\% \quad \text{daraus resultiert} \quad H_1: p < p_0 = 40\%$$

Das Gleichheitszeichen gehört immer zu  $H_0$ .

Wenn er recht hat, oder die Wahrscheinlichkeit noch größer ist, als 40%, dann interessiert uns das als Gegenpartei nicht. In diesen Fällen ist er sogar besser als er selbst angibt, was für uns zum Vorteil ist. Uns interessiert, ob er schlechter ist, als er uns einreden will. Deshalb interessiert uns der linke Teil der Verteilungskurve.

Merke : Die Bedingung von  $H_1$  zeigt den Hypothesentest an, der gemacht werden muss. Bei  $H_1$  steht

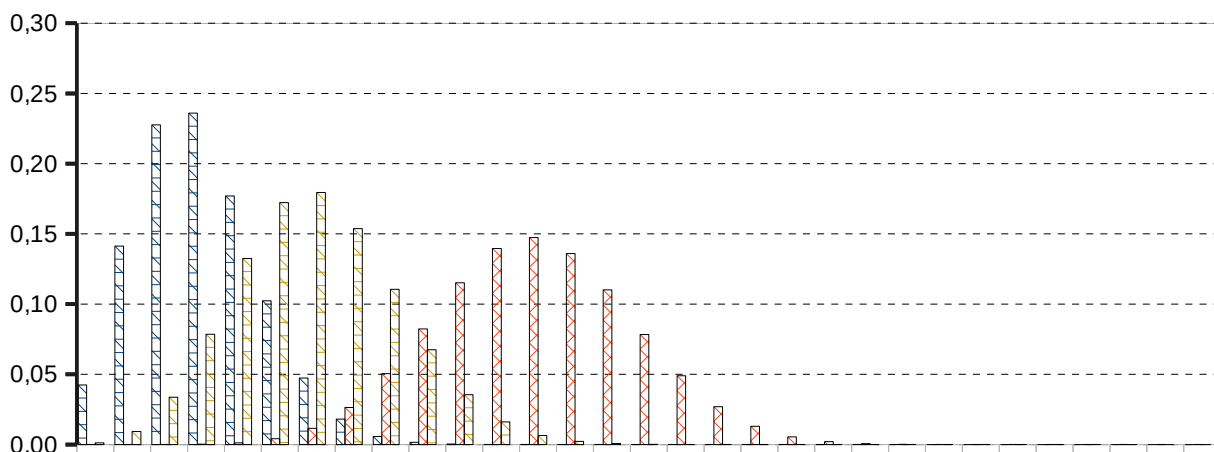
$$p < p_0$$

Die Spitze des Ungleichheitszeichen zeigt nach links, oder zu  $p$  Werten, die kleiner als  $p_0$  sind.

In diesem Fall handelt es sich um einen linksseitigen Test.

Uns interessiert, ob die Ereignisse nicht doch seltener auftreten, als bei 40% angenommen. Bei  $p = 0,4$  und  $n = 30$  erhält man einen Erwartungswert von 12, der auch die höchste Säule in den roten Säulen darstellt. Wenn man alle 30 Exemplare untersucht, sollte man 12 Stück finden, die die Eigenschaft erfüllen. Natürlich wird das nicht immer der Fall sein. Maschinen machen Fehler nicht nach mathematischen Regeln. Finden wird mehr, kann uns das nur recht sein. Finden wird weniger, sollte man irgendwann mißtrauisch werden.

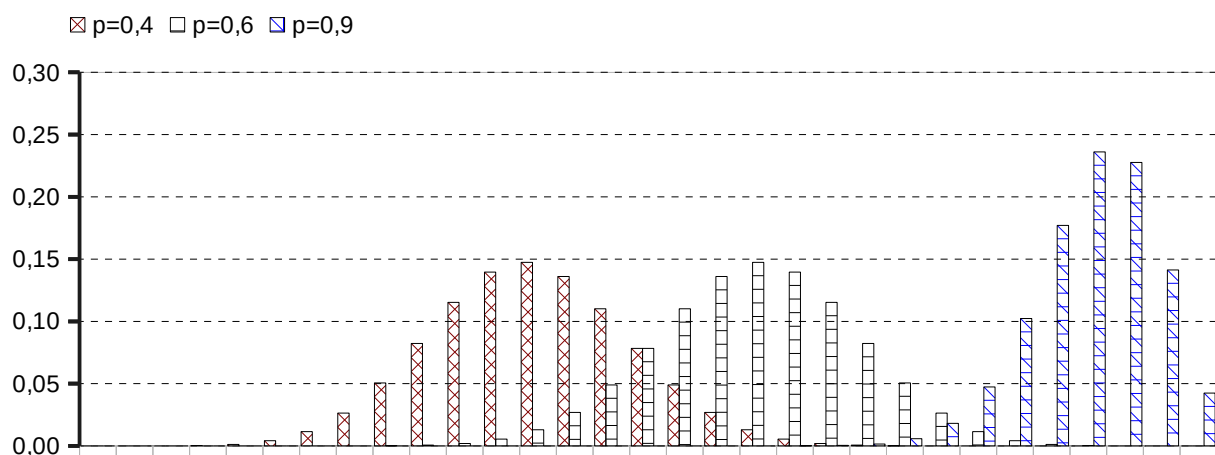
□  $p=0,1$    □  $p=0,4$    □  $p=0,2$    □  $p=1$



Finden wir nur 9 oder 10 Teile mit der gewollten Eigenschaft kann das noch in Ordnung sein, aber bei nur 3 oder 4 Teilen muss man an den 40% eher zweifeln. Bei den roten Säulen tritt  $k = 3$  oder  $k = 4$  mit einer sehr geringen Wahrscheinlichkeit auf, deshalb wäre es ein großer Zufall, wenn man tatsächlich einen solch extremen Test erhält. In diesem Fall ist wohl eher anzunehmen, dass die Wahrscheinlichkeit in Richtung der gelben und blauen Säulen geht.

### 34.2.3.1 $H_0$ „mindestens“ und $H_1$ „weniger als“ ist ein linksseitiger Test

Um „weniger als 40%“ zu bestätigen muss die Hypothese  $H_0$  so aufbauen, dass sie widerlegt wird, dh man muss die Hypothese  $H_0$  mit „mindestens“ erstellen. Im nachfolgenden Diagramm sind verschiedene Wahrscheinlichkeiten der Hypothese „mindestens 40%“ dargestellt.



Eine Hypothese „weniger als 40%“ kann man nur bestätigen, wenn man die  $H_0$  Hypothese „mindestens 40%“ ablehnt. Damit wird „weniger als 40%“ zur Hypothese  $H_1$ , die dann gilt, wenn man die Hypothese  $H_0$  ablehnen kann. Erreicht man dabei einen  $k$  Wert von 4,5 oder 6 muss man die Hypothese „mindestens 40%“ ablehnen und bestätigt damit die Hypothese „weniger als 40%“.

Die Hypothese „weniger als 40%“ wird somit zum linksseitigen Hypothesentest. Gegenhypothese  $H_1$  zur  $H_0$  Hypothese „mindestens 40%“. Damit ist „mindestens 40%“ ebenfalls ein linksseitiger Hypothesentest.

Wenn in der Aufgabe das Wort **mindestens** ( $p \geq p_0$ ) auftaucht, dann wissen wir, dass das die  $H_0$ -Hypothese kennzeichnet und wir einen **linksseitigen Hypothesentest** durchführen.

Wenn in der Aufgabe das Wort **weniger als** oder auch **weniger** ( $p < p_0$ ) auftaucht, dann wissen wir, dass das die  $H_1$  - Hypothese kennzeichnet und wir einen **linksseitigen Hypothesentest** durchführen.

„mindestens“ enthält das Gleichheitszeichen, da die Angabe als untere Grenze gemacht ist, und ist damit  $H_0$ ;

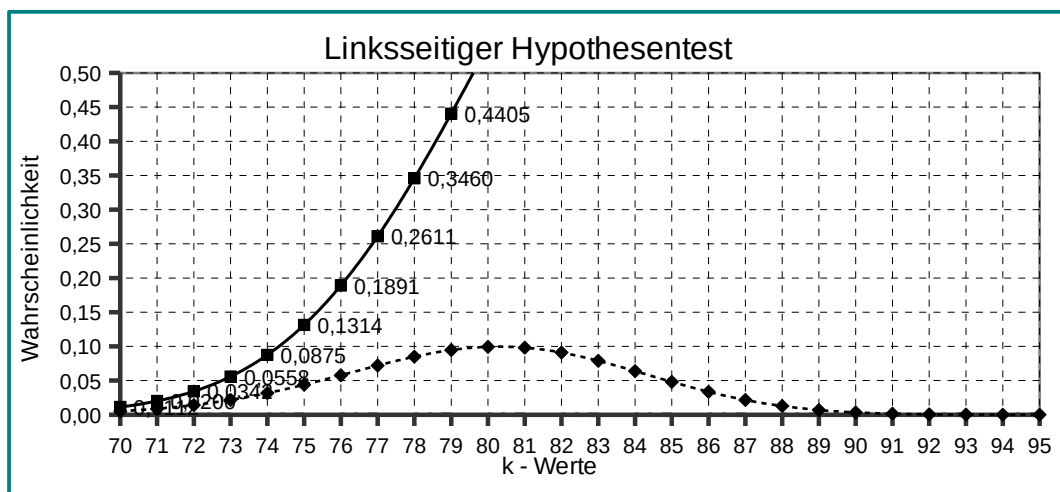
„weniger als“ enthält kein Gleichheitszeichen und gehört deshalb zu  $H_1$ ,

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

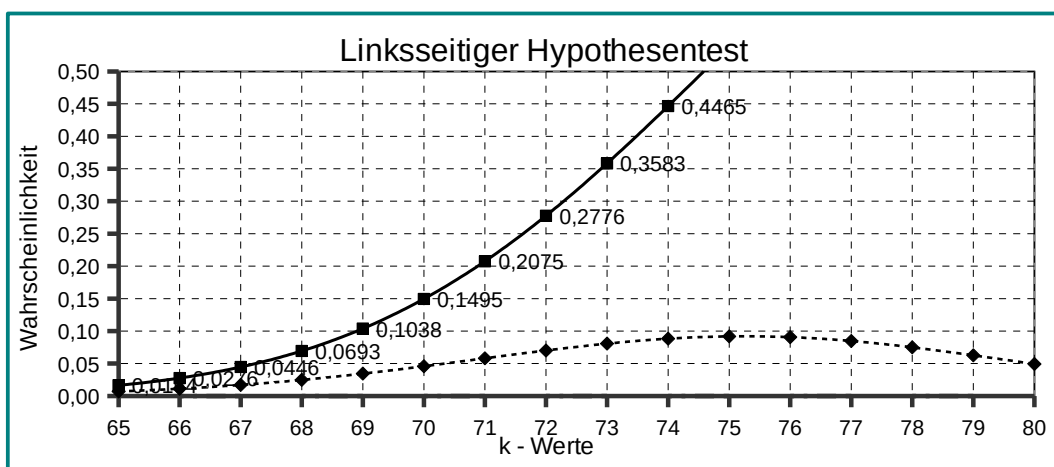
Der Anbieter eines Medikamentes sagt: Sein Produkt hat eine Heilungsquote von 80% ( $p = 80\%$ ). Wir als Nutzer haben Zweifel an der Quote und meinen er übertreibt:  $p < 80\%$ . **Das ist die Hypothese  $H_1$ .** Wenn die Heilungsquote größer ist, haben wir Pech gehabt, also gehören alle Wahrscheinlichkeiten  $p > p_0$  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:  $p < 80\%$ .

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  $\alpha = 0,1$



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 %. Beim Hypothesentest geht es aber um die summierte Wahrscheinlichkeit und das ist die obere Funktion. Dort ist der Wert für  $k = 74$ : 0,0875 und für  $k = 75$  beträgt der Wert 0,1314, damit größer als 0,1 und somit im Annahmehereich.



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 Annahmehbereich vom Erwartungswert aus so weit nach links verschoben, dass die Möglichkeit eines Auftretens noch kleinerer  $k$  - Werte als selten bei der angegebenen Wahrscheinlichkeit ansehen werden muss.

Im angegebenen Fall liegt der Erwartungswert bei  $E(X) = n \cdot p = 100 \cdot 0.08 = 80$ . Man räumt die Möglichkeit ein, dass man zufällig eine ungünstige Probe zieht und lässt noch zu, dass weniger Personen von dem Medikament geheilt werden, aber noch weniger als 74 lassen Zweifel aufkommen.

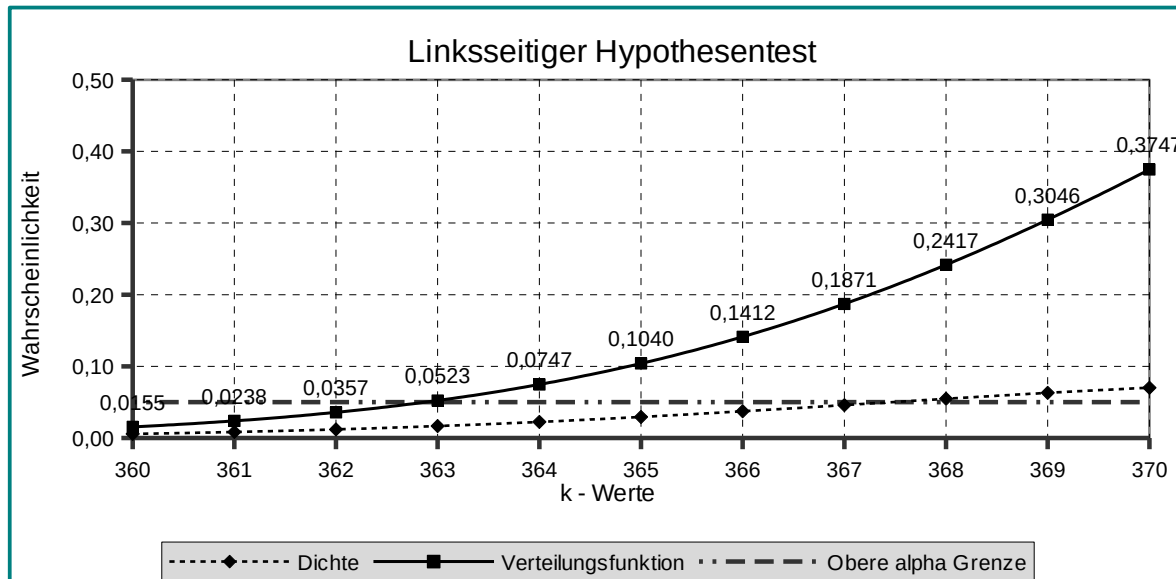
Beispiel.

Ein Großhändler verkauft Glühbirnen an einen Baumarkt. Er gibt an, daß von den Glühbirnen (mindestens) 93% 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 Entscheidungsregel lauten, so dass der Baumarkt in 95% der Fälle richtig liegt ?

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. Größere Wahrscheinlichkeiten verschieben die Funktion nach rechts. Das wollen wir nicht, das ist für uns ungünstig, da es die Behauptung des Herstellers bestätigt. 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  $\alpha$  (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.

Folgendes Vorgehen ist empfehlenswert:

#### Ablaufschema beim linksseitigen Hypothesentest

1. Festlegung der Hypothesen:  $H_0: p \geq p_0; H_1: p < p_0$
2. Festlegung des Stichprobenumfangs  $n$  und der Irrtumswahrscheinlichkeit  $\alpha$ ;
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  $g_L$  aus der Bedingung  $P(X \leq g_L) \leq \alpha$  und damit des Ablehnungsbereichs  $K = \{0, \dots, g_L\}$

4.1 ) Ablehnungsbereich bekannt	4.2 ) Irrtumswahrscheinlichkeit bekannt
Bestimmung der Irrtumswahrscheinlichkeit $\alpha$ bei gegebenem Ablehnungsbereich $P(X \leq g_L) \leq \alpha$	Bestimmung der linken Grenze $g_L$ aus der gegebenen Irrtumswahrscheinlichkeit $\alpha$ : $P(X \leq g_L) \leq \alpha$ und damit des Ablehnungsbereichs $K = \{0, \dots, g_L\}$

5. Angabe der Entscheidungsregel oder Entscheidung aufgrund eines konkreten Stichprobenergebnisses

Bei einem linksseitigen Hypothesentest ist immer eine Wahrscheinlichkeit als untere Grenze zu prüfen. Gesucht ist in der Binomialverteilung dasjenige  $k$ , für das die **summierte Wahrscheinlichkeit** kleiner ist als der vorgegebene Grenzwert  $\alpha$  ist. Es ist also immer mit der Verteilungsfunktion  $F_{n;p}$  zu arbeiten und nicht mit  $B_{n;p}$ .

### 34.2.3.3 Linksseitiger Hypothesentest : Irrtumswahrscheinlichkeit gesucht

Die Anzahl der akzeptierten Ereignisse ist festgelegt, wie groß ist die Wahrscheinlichkeit, dass wir uns irren An der Stelle kommt die Irrtumswahrscheinlichkeit ins Spiel.

Wir akzeptieren seine Behauptung, wenn wir wenigstens noch 363 Teile mit der geforderten Eigenschaft finden. Sind es weniger meinen wir, seine Angaben stimmen nicht. Gleichzeitig sehen wir an der Linie, dass bei einer Wahrscheinlichkeit von 0.95 durchaus die Ereignisse mit  $k=362$  oder  $k=361$  auftreten können, und damit die Behauptung doch richtig ist, obwohl wir sie nicht akzeptieren. .

Wenn wir  $k=363$  als Grenze festlegen, dann kann es sein, dass wir uns mit einer Wahrscheinlichkeit  $P(X < 363)$  irren, also der Summe aller möglichen Ereignisse deren  $k$  kleiner als 363 ist. Das ist die Irrtumswahrscheinlichkeit bei  $k=363$ .

Für jeden Wert  $k$  erhält man eine andere Irrtumswahrscheinlichkeit.

Die Berechnung dieser Irrtumswahrscheinlichkeit ist sehr einfach, man muss nur die Funktion Binomcdf () benutzen  $n (=400)$  liegt fest,  $p (=0,95)$  liegt fest und auch  $k (=363)$  liegt fest. Die berechnete Wahrscheinlichkeit ist die gesuchte Irrtumswahrscheinlichkeit.

$$\text{Binomcdf}(400; 0,95; 363)$$

Es wurde als Grenze die Stückzahl 363 festgelegt. Damit muß die daraus resultierende Wahrscheinlichkeit als Irrtumswahrscheinlichkeit akzeptiert werden.

### 34.2.3.4 Linksseitiger Hypothesentest : Ablehnungsbereich gesucht

Hier geht es um den umgekehrten Fall. Die Wahrscheinlichkeit, die ich als Irrtum zulasse, ist festgelegt, welches  $k$  kann ich dann noch akzeptieren. Ich weiß, dass ich mich auch irren kann, möchte aber diesen Irrtum so gering, wie möglich halten. z.B kann man sagen : Ich akzeptiere, dass ich mich in 5% aller Fälle auch irren kann, welche (geringere) Anzahl von positiven Ereignissen kann ich dann noch akzeptieren. In der Formulierung der Wahrscheinlichkeitsrechnung heißt das:

Für welches  $k_0$  ist die summierte Wahrscheinlichkeit von 0 bis  $k_0$  noch kleiner als die vorgegebenen 5% :  $P_{400; 0,95}(X \leq k_0) < 0,05$

Auf Grund der Komplexität der Formel der Binomialverteilung lässt sich dieser Wert nicht direkt berechnen, da die summierte Binomialverteilung nicht nach  $k$  aufgelöst werden kann. Die Berechnung ist nur mittels Listen oder dem Umstieg auf die Normalverteilung möglich.möglich.

#### **Linksseitiger Hypothesentest:**

- Gesucht ist das  $k$ , bei dem die Wahrscheinlichkeit von 5% das erste mal überschritten wird.
- dieses  $k$  ist der Beginn des Annahmebereichs.
- $k - 1$  ist das Ende des Ablehnungsbereichs.

## (1) Berechnung über Listen

Wo beginnt man mit der Liste. Gesucht ist ein linksseitiger Test mit einer Irrtumswahrscheinlichkeit von 5%. Nach den Sigmaregeln ist das ein einseitiges Sigmaintervall von 95%. 95% entsprechen einem Faktor von 1,64, also beginnt man mit 1,5.

$$372 - 1,5 \cdot 5,1 = 364,3$$

$$P(X \leq 362) \approx 0,0357$$

$$P(X \leq 363) \approx 0,0522$$

$$P(X \leq 364) \approx 0,0746$$

Beim linksseitigen Hypothesentest ist das erste  $k$ , das die Irrtumswahrscheinlichkeit überschreitet das erste  $k$  des Annahmebereichs.  
Der Ablehnungsbereich endet bei  $k-1$ .

Der gesuchte Ablehnungsbereich ist  $\bar{A} = \{ 0; \dots; 362 \}$ .

Es zeigt sich auch hier, daß das Signifikanzniveau von 5% bei  $k = 363$  überschritten wird, damit ist  $k = 363$  der Beginn des Annahmebereichs.

X	Y1
362	.03571
363	.05228
364	.07466
365	.10397
366	.14122
367	.18707
368	.24169
Y1 = .052281226573	

## (2) Berechnung über Quantile der Normalverteilung

$$p = 93\%$$

$$n = 400$$

$$\alpha = 5\%$$

$$\mu = n \cdot p = 400 \cdot 0,93 = 372$$

$$\sigma = \sqrt{n \cdot p \cdot q}$$

$$= \sqrt{400 \cdot 0,93 \cdot 0,07} = 5,1$$

$$\text{InvNorm}(\alpha; \mu; \sigma) = k$$

$$\text{InvNorm}(0,05; 372; 5,1) = 363,6$$

In die Inverse Normalverteilung sind die Werte für  $p$ ,  $\mu$  und  $\sigma$  einzugeben. Das Ergebnis ist ein  $k$  Wert, bei dem die Irrtumswahrscheinlichkeit das erste Mal überschritten wird.

Da geringfügige Abweichungen zur Binomialverteilung auftreten können, sind die  $k$  Werte in der Umgebung des berechneten Wertes noch mal genau zu untersuchen.

(Die Reihenfolge der Eingabe der Werte in den Taschenrechner ist von Rechner zu Rechner verschieden, entweder im Buch nachschlagen oder ausprobieren, eventuell ist auch eine Eingabemaske vorhanden)

Linksseitiger Test:

$$\text{Ablehnungsbereich bis } 362 : \text{binomcdf}(400; 0,93; 362) = 0,0357$$

$$\text{Annahmebereich ab } 363 : \text{binomcdf}(400; 0,93; 363) = 0,0523$$

Da hier die Verteilungsfunktion gewechselt wurde ist es empfehlenswert, dass man den  $k$  Wert mit der Binomialverteilung nachkontrolliert.

## (3) Berechnung über Quantile der Standard Normalverteilung

Für die Standard Normalverteilung  $\mu = 0$  und  $\sigma = 1$  belassen.

$$\text{InvNorm}(\alpha ; 0 ; 1) = k$$

$$\text{InvNorm}(0,05 ; 0 ; 1) = -1,6448$$

Der angegebene Wert  $-1,6448$  ist der Faktor  $c$  für die Größe des Sigma Intervalls.

Bei der Inversen Standard Normalverteilung lässt man  $\mu = 0$  und  $\sigma = 1$ . Das Ergebnis ist nicht der gesuchte  $k$  Wert, sondern der Faktor  $c$  für das Sigma Intervall. Ein Signifikanzniveau von 2% für einen rechtsseitigen Test führt zu  $c = \Phi^{-1}(0,98) = 2,0537$ .

$$\begin{aligned} \mu &= 372 \\ \sigma &= 5,1 \end{aligned}$$

$$k_0 = \mu + c \sigma = 372 - 1,6448 \cdot 5,1 = 363,6$$

Nach dieser Berechnung ist  $k = 363$  das Ende des Ablehnungsbereichs und  $k = 364$  der Beginn des Annahmebereichs, da die Irrtumswahrscheinlichkeit von 0,05 genau zwischen diesen beiden Ganzzahlen erreicht wird. Über Binomialverteilung nachprüfen.

## (4) Berechnung über Inverse Binomialverteilung

Manche Rechner besitzen auch eine Inverse Binomialverteilung. Über diese lässt sich der gesuchte  $k$  Wert ebenfalls berechnen.

$$\text{InvBinomcdf}(n ; p ; \alpha) = k$$

$$\text{InvBinomcdf}(400 ; 0,93 ; 0,05) = 363$$

363 ist der erste  $k$  Wert, bei dem die Irrtumswahrscheinlichkeit überschritten wird. Bei einem linksseitigen Test bedeutet das, der Beginn des Annahmebereichs. Damit endet der Ablehnungsbereich bei  $k = 362$ .

(Auch Taschenrechner können die Binomialverteilung nicht nach  $k$  auflösen. Sie summieren von 0 an so lange, bis sie die vorgegebene Wahrscheinlichkeit  $\alpha$  überschritten haben. Dazu kommt noch, daß es in der Binomialverteilung nicht möglich ist, die 5% genau zu treffen, da keine reellen Werte für  $k$  möglich sind. Deshalb gibt der Taschenrechner immer das  $k$  zurück, das das vorgegebene  $\alpha$  als erstes überschritten hat.)

### 34.2.4 Rechtsseitiger Hypothesentest mit Binomialverteilung

Bei einem rechtsseitigen Hypothesentest sprechen (zu)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.

Jemand behauptet, sein Ereignis (defekte Teile, faule Äpfel) hat eine Wahrscheinlichkeit von **höchstens 40%**:  $H_0: p \leq p_0 = 40\%$  daraus resultiert  $H_1: p > p_0$   
Das Gleichheitszeichen gehört immer zu  $H_0$ .

Wenn er recht hat, oder die Wahrscheinlichkeit noch kleiner ist, als 40%, dann interessiert uns das als Gegenpartei nicht. In diesen Fällen ist er sogar besser als er selbst angibt, was für uns zum Vorteil ist. Uns interessiert, ob er schlechter ist, als er uns einreden will. Deshalb interessiert uns der rechte Teil der Verteilungskurve.

Merke : Die Bedingung von  $H_1$  zeigt den Hypothesentest an, der gemacht werden muss. Bei  $H_1$  steht  

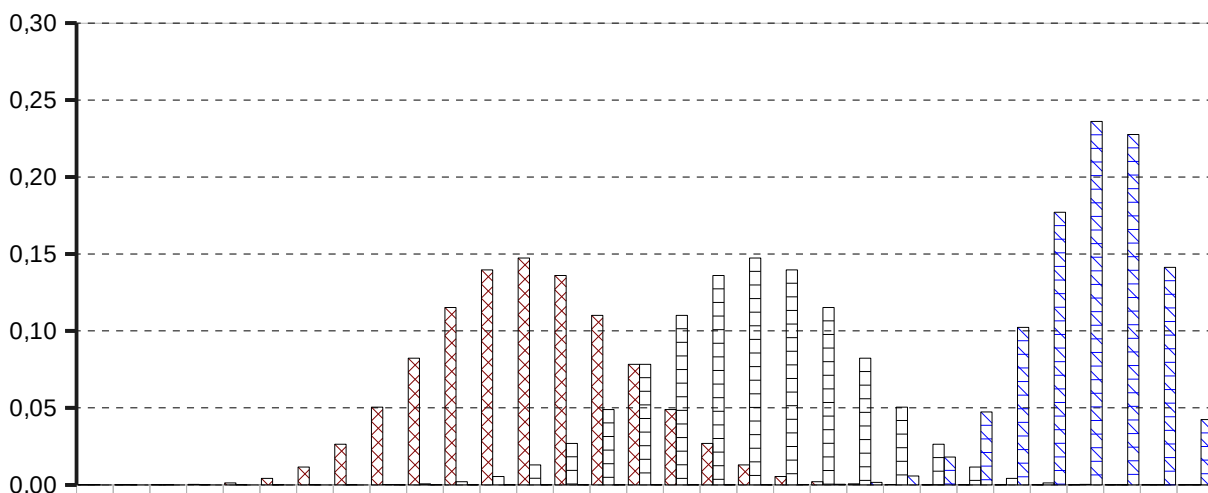
$$p > p_0$$
 Die Spitze des Ungleichheitszeichen zeigt nach rechts, oder zu  $p$  Werten, die größer als  $p_0$  sind.  
 In diesem Fall handelt es sich um einen rechtsseitigen Test.

Für uns als Gegenpartei eher interessant die Werte  $k = 17, 18, 19$ . Bei einer Wahrscheinlichkeit von 40% dürften diese Werte eher selten auftreten. Da sie aber dennoch auftreten können, unterliegen wir auch in diesen Fällen einer Irrtumswahrscheinlichkeit, und zwar der Summe der Wahrscheinlichkeiten von 18, 19, 20 bis zu  $n (= 30)$ .

#### 34.2.4.1 $H_0$ „höchstens“ und $H_1$ „mehr als“ ist ein rechtsseitiger Test

Die Schwierigkeit der Unterscheidung liegt z.B im Unterschied zwischen „höchstens“ und „weniger als“. Man könnte meinen, das ist das gleiche, ist es aber nicht ! Die Zusammenhänge sollen hier wieder an  $n = 30$  demonstriert werden.

☒  $p=0,4$    ☐  $p=0,6$    ☒  $p=0,9$



Aus dem Bild ist zu erkennen, dass ab etwa  $k = 20$  die folgenden  $k$  Werte von 21 bis 30 kaum noch Wahrscheinlichkeiten besitzen, und damit sehr selten auftreten müssen. Bezüglich des Hypothesentest bedeutet das, dass der Ablehnungsbereich rechts liegt, und damit ein rechtsseitiger Test vorliegt.

Erzielt man ein  $k$ , das größer als 18 oder 19 ist, hat man „höchstens 40%“ widerlegt und damit als  $H_1$  die Hypothese „mehr als 40%“ bestätigt. Damit ist „mehr als“ ein rechtsseitiger Test.

Erzielt man ein  $k$ , das kleiner als 17 ist, hat man „höchstens“ nicht widerlegt, aber auch nicht bestätigt, da man Hypothesen nicht bestätigen kann. Demzufolge hat man damit auch nicht „weniger als“ bestätigt.

Wenn in der Aufgabe das Wort **höchstens** ( $p \leq p_0$ ) auftaucht, dann wissen wir, dass das die  $H_0$ -Hypothese kennzeichnet und wir einen **rechtsseitigen Hypothesentest** durchführen.

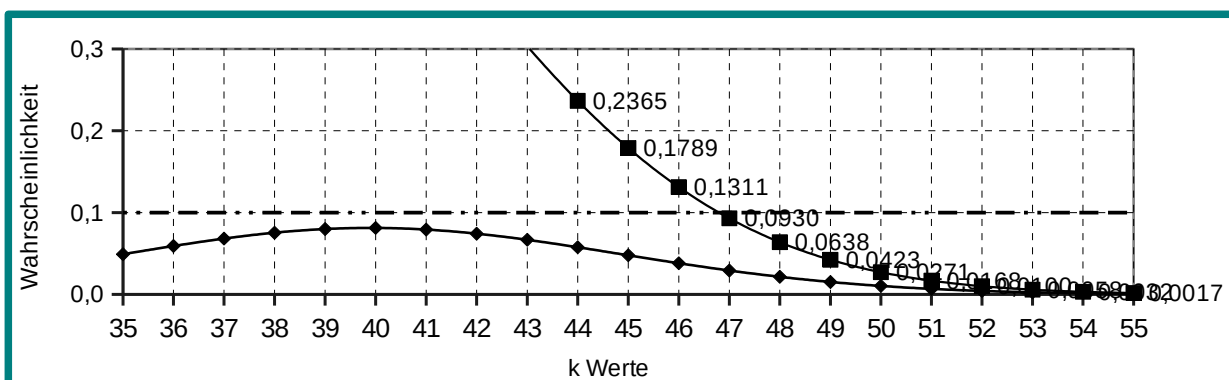
Wenn in der Aufgabe das Wort **mehr als** oder auch **größer** ( $p > p_0$ ) auftaucht, dann wissen wir, dass das die  $H_1$  - Hypothese kennzeichnet und wir einen **rechtsseitigen Hypothesentest** durchführen.

„höchstens“ enthält das Gleichheitszeichen, da die Angabe als obere Grenze gemacht ist, „mehr als“ enthält kein Gleichheitszeichen. Damit „höchstens“ zu  $H_0$ ; und „mehr als“ zu  $H_1$ .

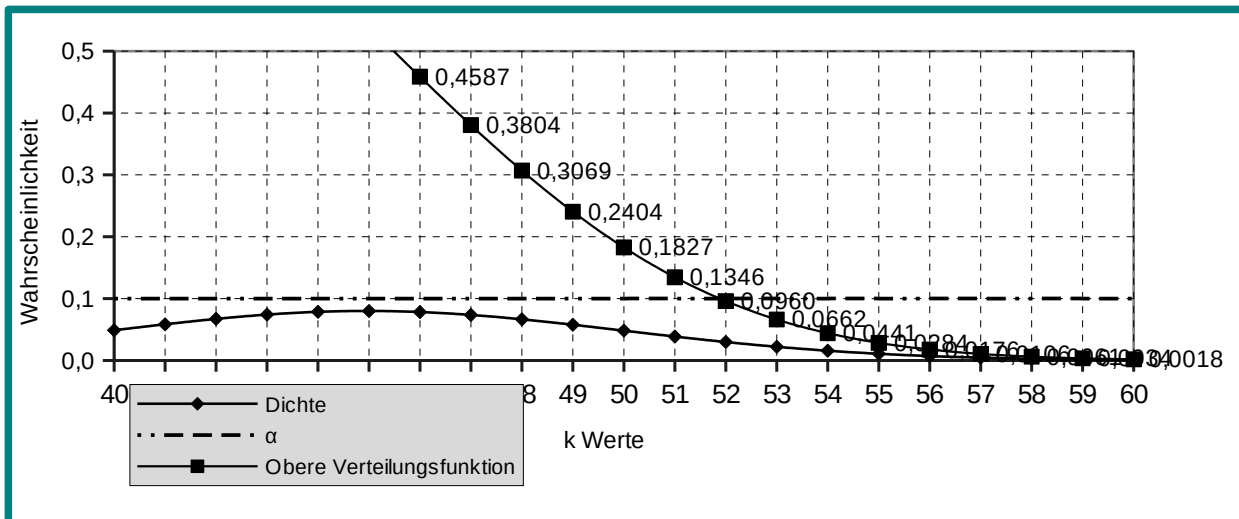
#### 34.2.4.2 Warum steht beim rechtsseitigen Hypothesentest $H_0 : p \leq p_0$

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  $\alpha = 0,1$



Die Grenze des Ablehnungsbereichs für  $p = 0,4$  liegt bei  $k = 47$ . Der Wert 47 selbst tritt nur mit einer Wahrscheinlichkeit 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. (Die obere Kurve ist die summierte Wahrscheinlichkeit.)



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

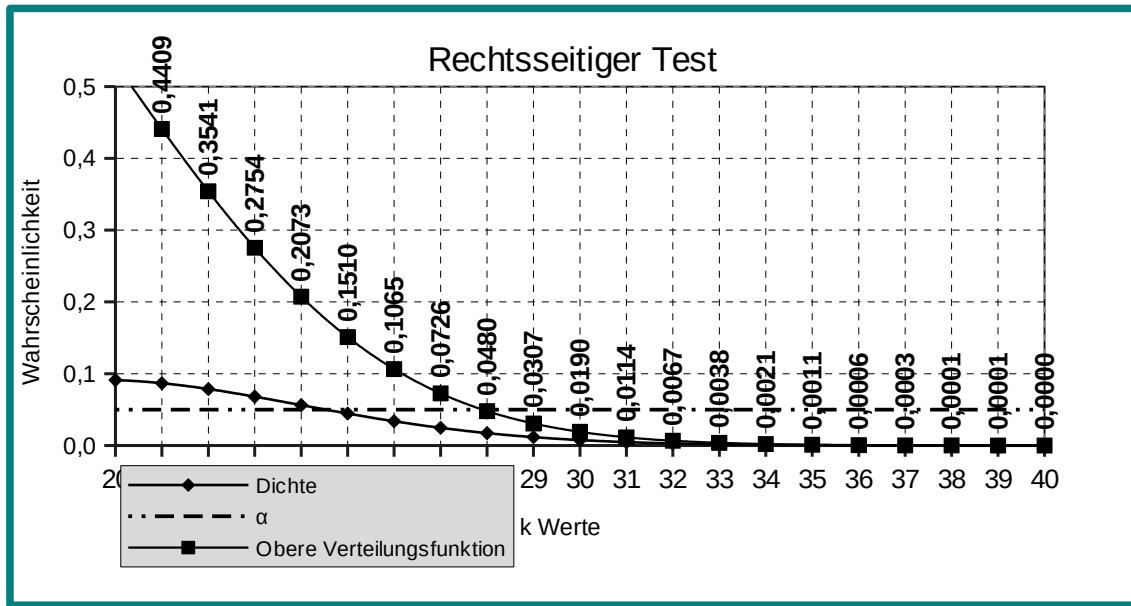
- Ablehnungsbereich so weit nach rechts verschoben, dass man bei noch größeren Werten eher Zweifel an der Wahrscheinlichkeit hat, als dass man eine ungünstige Stichprobe erhalten hat.
- der Annahmebereich bereits so weit nach rechts verschoben ist dass man bei noch größeren Werten an der Wahrscheinlichkeit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

Im angegebenen Fall liegt der Erwartungswert bei  $E(X) = n \cdot p = 100 \cdot 0,04 = 40$ . 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.

Zur weiteren Darstellung soll hier wieder das Beispiel mit den Glühbirnen aus dem Linksseitigen Test verwendet werden, aber mit einer „höchstens“ Bedingung von dem Lieferanten.

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 Entscheidungsregel lauten, so dass der Baumarkt in 95% der Fälle richtig liegt ?





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 5% sollten bei einer „normalen“ Stichprobe etwa 19 oder 20 defekte Glühbirnen auftreten. Der Erwartungswert liegt bei  $E(X) = n \cdot p = 400 \cdot 0,05 = 20$ . 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 unwahrscheinlich. 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

Folgendes Vorgehen ist empfehlenswert:

**Ablaufschema beim rechtsseitigen Hypothesentest**

1. Festlegung der Hypothesen:  $H_0: p \leq p_0; H_1: p > p_0$
2. Festlegung des Stichprobenumfangs  $n$  und der Irrtumswahrscheinlichkeit  $\alpha$
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  $K = \{g_R, \dots, n\}$

4.1 ) Ablehnungsbereich bekannt	4.2 ) Irrtumswahrscheinlichkeit bekannt
Bestimmung der Irrtumswahrscheinlichkeit $\alpha$ bei gegebenem Ablehnungsbereich $P(X \geq g_R) \leq \alpha$	Bestimmung der rechten Grenze $g_R$ aus der gegebenen Irrtumswahrscheinlichkeit $\alpha$ : $P(X \geq g_R) \leq \alpha$ und damit des Ablehnungsbereichs $K = \{ g_R, \dots, n \}$

5. Angabe der Entscheidungsregel oder Entscheidung aufgrund eines konkreten Stichprobenergebnisses

### 34.2.4.3 Rechtsseitiger Hypothesentest : Irrtumswahrscheinlichkeit gesucht

Die Anzahl der akzeptierten Ereignisse ist festgelegt, wie groß ist die Wahrscheinlichkeit, dass wir uns irren An der Stelle kommt die Irrtumswahrscheinlichkeit ins Spiel.

Wir akzeptieren seine Behauptung, wenn wir höchstens 27 fehlerhafte Teile finden. Sind es mehr meinen wir, seine Angaben stimmen nicht. Gleichzeitig sehen wir an der Linie, dass bei einer Irrtumswahrscheinlichkeit von 0.05 durchaus die Ereignisse mit  $k = 28$  oder  $k = 29$  auftreten können, und damit die Behauptung doch richtig ist, obwohl wir sie nicht akzeptieren. .

Wenn wir  $k = 27$  als Grenze festlegen, dann kann es sein, dass wir uns mit einer Wahrscheinlichkeit  $P(X > 27)$  irren, also der Summe aller möglichen Ereignisse deren  $k$  größer als 27 ist. Das ist die Irrtumswahrscheinlichkeit bei  $k = 27$ .

Wir wollen die Behauptung noch akzeptieren, wenn die gefundenen defekten Teile höchstens 27 ist. Damit ist unsere Irrtumswahrscheinlichkeit  $P(X > 27) = P(X \geq 28)$ .

Leider lässt sich die Wahrscheinlichkeit für  $\geq$  nicht berechnen, da der GTR nur Berechnungen für  $\leq$  zur Verfügung stellt. Deshalb müssen wir hier die Wahrscheinlichkeit für

$$P(X \geq 28) = 1 - P(X \leq 27)$$

berechnen. Das kann ohne Probleme mit dem GTR erledigt werden.

Des Ende unseres Annahmebereichs ist 27 und der Beginn des Ablehnungsbereichs ist 28. Nach der obigen Grafik ist die Irrtumswahrscheinlichkeit dann 0,048 , der oberen Summenwahrscheinlichkeit ab  $k = 28$ .

### 34.2.4.4 Rechtsseitiger Hypothesentest : Ablehnungsbereich gesucht

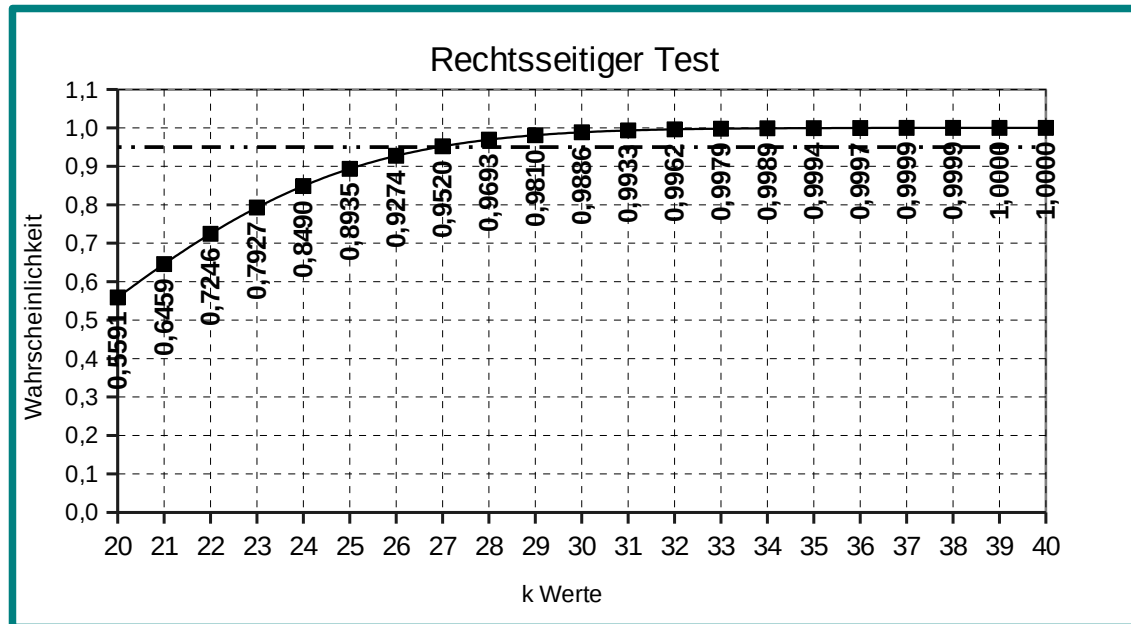
Für welches  $k$  ist die summierte Wahrscheinlichkeit von  $k$  bis  $n$  noch kleiner als die vorgegebenen 5% :  $P_{400; 0,05}(X \geq k_0) < 0,05$   
 Oder  
 Bis zu welchem  $k$  muss summiert werden, damit die Summe von 0 bis  $k_0 - 1$  größer als 0,95 ist :  $P_{400; 0,05}(X \leq k_0 - 1) > 0,95$

$k_0$  ist der Beginn des Ablehnungsbereichs. Bestimmen lässt sich aber nur  $k_0 - 1$ .

Diese Berechnung wird wieder etwas schwieriger, wie beim Linksseitigen Test. Die Berechnung erfolgt meist über Listen. Dazu kommt noch, daß eine Summierung bis  $n$  durch die Funktionen im Rechner nicht angeboten werden und auf eine Berechnung von 0 bis  $k$  umgerechnet werden muß. Wie man diesen Problemen etwas entgegen kann, wird am Ende diese Kapitel an einem Beispiel gezeigt.

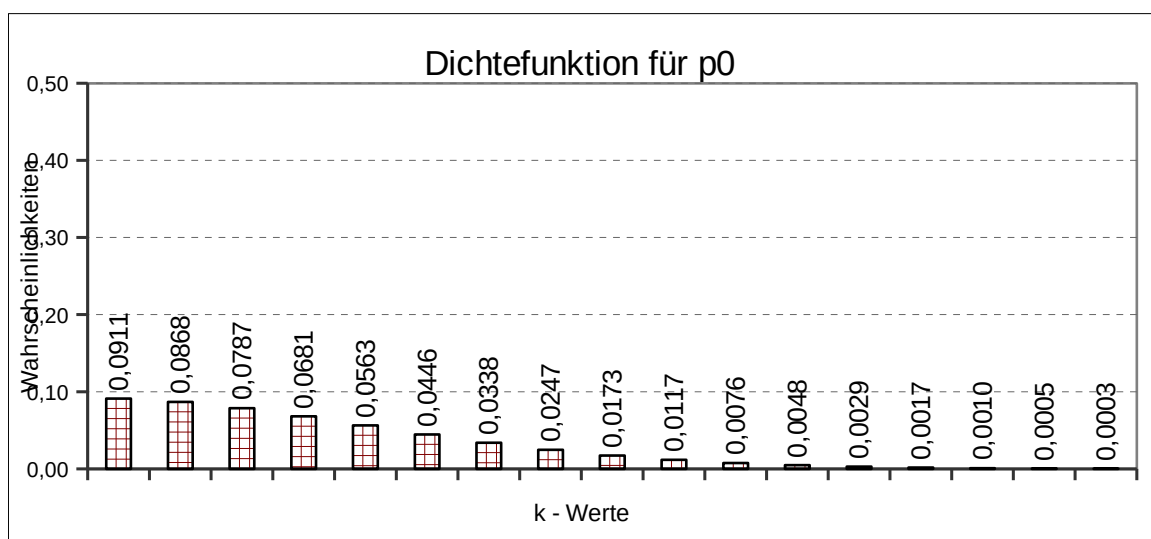
Zunächst sollen die Überlegungen für die Berechnung der Grenze zwischen Annahme- und Ablehnungsbereich allgemein untersucht werden.

Es wird dasjenige  $k$  gesucht, bei dem die summierte Wahrscheinlichkeit das erste Mal die 95% überschreitet. Das ist aber nicht das  $k$ , welches den Beginn des Ablehnungsbereiches angibt, sondern auf grund der Umrechnung das um 1 niedrigere, damit das  $k$ , das das Ende des Annahmebereichs angibt. Dazu soll die folgende summierte Wahrscheinlichkeitsfunktion für  $p = 0,05$  betrachtet werden.



Aus der Kurve der summierten Wahrscheinlichkeitsfunktion ist zu erkennen, dass das bei  $k = 27$  das erste Mal die 95% überschritten werden. Daraus wird oft für den rechtsseitigen Test die falsche Schlußfolgerung gezogen, damit bei  $k = 27$  der Ablehnungsbereich beginnt. Dieser Wert stellt beim rechtsseitigen Test das Ende des Annahmebereichs dar. Der Ablehnungsbereich beginnt bei dem nächsten  $k$ .

Dazu betrachten wir für  $n = 400$  und  $p = 0,05$  die Einzelwahrscheinlichkeiten. Alle Wahrscheinlichkeiten von  $k = 0$  bis  $k = 27$  ergeben zusammen  $0,9520$ , wie aus der oberen Darstellung zu entnehmen ist. Damit ist die Summe von  $27$  bis  $400 = 1 - 0,9520 = 0,0480$ .



Jetzt sollen die Einzelwahrscheinlichkeiten von  $n$  rückwärts bis zu  $k = 27$  betrachtet werden. Aus der unteren Darstellung der Einzelwahrscheinlichkeiten sieht man, dass die Wahrscheinlichkeiten von 35 bis 400 so klein sind, dass sie erst ab der 4. Stelle nach dem Komma verwertbare Zahlen liefern. Deshalb lassen wir sie hier weg, weil sie sehr klein sind. Wir beginnen mit der Betrachtung bei  $k = 34$ .

K	34	33	32	31	30	29	28	27
$P(X=k)$	0,0010	0,0017	0,0029	0,0048	0,0076	0,0117	0,0173	0,0247
$P(k \leq X \leq n)$	0,0010	0,0027	0,0056	0,0103	0,0179	0,0296	0,0469	0,0715

Bis  $k = 28$  ist die Summe von  $n$  abwärts summiert noch kleiner als 0,05. Für  $k = 27$  gilt das aber nicht mehr. Dort liegt die Summe über 0,05. Damit gehört  $k = 27$  zum Annahmebereich und der Ablehnungsbereich beginnt mit  $k = 28$ .

### Rechtsseitiger Hypothesentest:

- Gesucht ist das  $k$ , bei dem die Wahrscheinlichkeit von 95% das erste mal überschritten wird.
- dieses  $k$  ist der Ende des Annahmebereichs.
- $k + 1$  ist der Beginn des Ablehnungsbereichs.

Die Kennzahlen des Tests ergeben folgende Werte

$$n = 400$$

$$p = 5\%$$

$$\alpha = 5\%$$

$$\mu = n \cdot p = 400 \cdot 0,05 = 20$$

$$\sigma = \sqrt{n \cdot p \cdot q} = \sqrt{400 \cdot 0,05 \cdot 0,95} = 4,36$$

#### (1) Berechnung über Listen

Wo beginnt man mit der Liste. Gesucht ist ein rechtsseitiger Test mit einer Irrtumswahrscheinlichkeit von 5%. Um die summierte Binomialverteilung nicht auf die unteren Summen umrechnen zu müssen, setzt man als Grenze nicht 5%, sondern 95% als Annahmebereich an.

Nach den Sigmaregeln ist das ein einseitiges Sigmaintervall von 95%. 95% entsprechen einem Faktor von 1,64, aber nicht mit einem Minus (linksseitiger Test) sondern mit einem Plus (rechtsseitiger Test) also beginnt man mit 1,5.

$$20 + 1,5 \cdot 4,36 = 26,52$$

$$P(X \leq 26) \approx 0,0357$$

$$P(X \leq 27) \approx 0,0522$$

$$P(X \leq 28) \approx 0,0746$$

Der gesuchte  
Ablehnungsbereich  
ist  $\bar{A} = \{28; \dots; 400\}$ .

X	$\Phi_1$
23	.79269
24	.84898
25	.89353
26	.92735
27	.95201
28	.9693
29	.98097

$X=27$

Beim rechtsseitigen Hypothesentest ist das erste  $k$ , das die Irrtumswahrscheinlichkeit unterschreitet das letzte  $k$  des Annahmebereichs. Der Ablehnungsbereich beginnt bei  $k+1$ .

## (2) Berechnung über Inverse Normalverteilung

$$\text{InvNorm}(\alpha, \mu, \sigma) = k$$

$$\text{InvNorm}(0,95, 20, 4,36) = 27,17$$

Rechtsseitiger Test:

Annahmereich bis 27 :

$$\text{binomcdf}(400 : 0.05 ; 27) = 0,952$$

Ablehnungsbereich ab 28 :

$$\text{binomcdf}\{400 : 0.05 : 28\} = 0.9693$$

In die Inverse Normalverteilung sind die Werte für  $p$ ,  $\mu$  und  $\sigma$  einzugeben. Das Ergebnis ist ein  $k$  Wert, bei dem die Irrtumswahrscheinlichkeit das erste Mal überschritten wird.

Da geringfügige Abweichungen zur Binomialverteilung auftreten können, sind die  $k$  Werte in der Umgebung des berechneten Wertes noch mal genau zu untersuchen.

Da hier die Verteilungsfunktion gewechselt wurde ist es empfehlenswert, dass man den  $k$  Wert mit der Binomialverteilung nachkontrolliert.

## (3) Berechnung über die Inverse Standard Normalverteilung

Für die Standard Normalverteilung  $\mu = 0$  und  $\sigma = 1$  belassen.

$$\text{InvNorm}(\alpha, \mu, \sigma) = c$$

$$\text{InvNorm}(0,95, 0, 1) = +1,6448$$

Der angegebene Wert +1,6448 ist der Faktor  $c$  für die Größe des Sigma Intervalls.

$$\mu + c \sigma = 20 + 1,6448 * 4,36 = 27,17$$

Bei der Inversen Standard Normalverteilung lässt man  $\mu = 0$  und  $\sigma = 1$ . Das Ergebnis ist nicht der gesuchte  $k$  Wert, sondern der Faktor  $c$  für das Sigma Intervall. Ein Signifikanzniveau von 2% für einen rechtsseitigen Test führt zu  $c = \Phi^{-1}(0,98) = 2,0537$ .

Nach dieser Berechnung ist  $k = 28$  das Ende des Annahmereichs und  $k = 29$  der Beginn des Ablehnungsbereichs, Da die Irrtumswahrscheinlichkeit von 0,05 genau zwischen diesen beiden Ganzzahlen erreicht wird. Über Binomialverteilung nachprüfen.

## (4) Berechnung über Inverse Binomialverteilung

Manche Rechner besitzen auch eine Inverse Binomialverteilung. Über die lässt sich der gesuchte  $k$  Wert berechnen.

$$\text{InvBinomcdf}(n, p, \alpha)$$

$$\text{InvBinomcdf}(400 ; 0,05 ; 0,95) = 27$$

27 ist der erste  $k$  Wert, bei dem die Irrtumswahrscheinlichkeit überschritten wird. Bei einem rechtsseitigen Test bedeutet das, das Ende des Annahmereichs. Damit beginnt der Ablehnungsbereich bei  $k = 28$ .