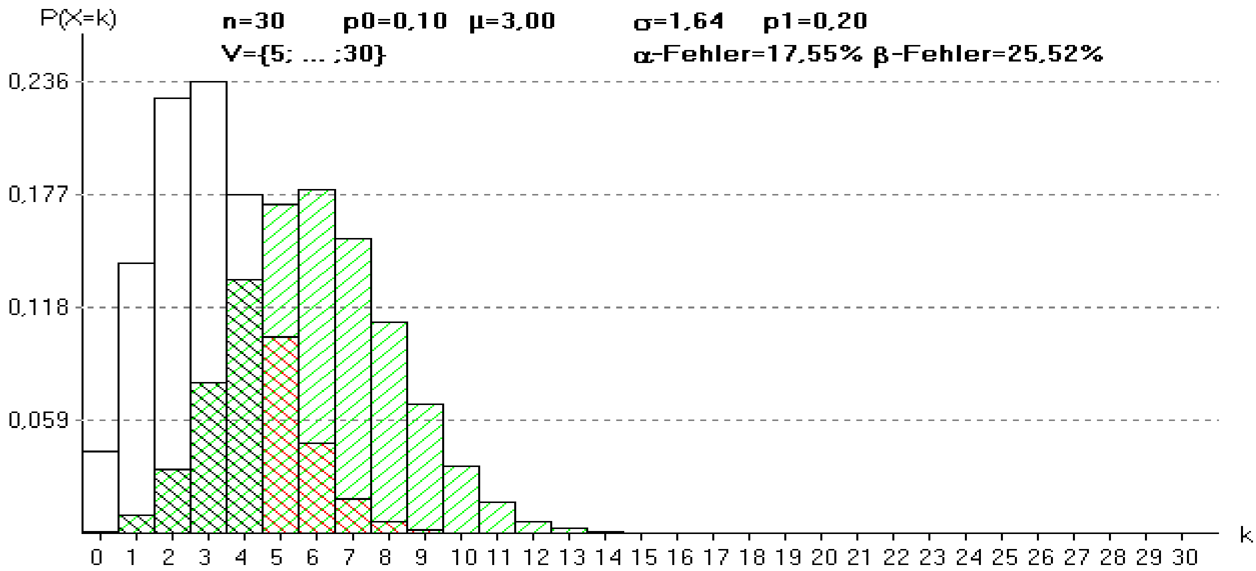
Alternativtest – Einführung und Aufgabenbeispiele Ac

Einführendes Beispiel:

Ein Medikament half bisher 10% aller Patienten. Von einem neuen Medikament behauptet der Hersteller, dass es 20% aller Patienten hilft. Diese Aussage wird an 30 Patienten getestet.
Wie lautet die Entscheidungsregel ? Wie groß sind die Fehler ?



Folgendes sei festgelegt (siehe Grafik oben):

n=30 H_o: p=0,1 H_1: p=0,2

Man sieht, dass die Grenze zwischen beiden Verteilungen offenbar zwischen k=4 und k=5 liegt.
Daher definieren wir als
Annahmebereich für H_o: A = {0; … ;4} und als
Verwerfungsbereich für H_o: V = {5; … ;30} .
Für die Gegenhypothese H_1 gilt natürlich das Umgekehrte !

Entscheidungsregel:
Entscheide dich für H_o, falls das Versuchsergebnis $X \leq 4$, andernfalls entscheide dich für H_1 .

Wir berechnen einige Wahrscheinlichkeiten:

$P_{0,1}(X \leq 4) = 82{,}4\%$ Also liegt das Versuchsergebnis mit 82,4%iger Wahrscheinlichkeit im Annahmebereich von H_o .
$P_{0,1}(X \geq 5) = 17{,}6\%$ Also liegt das Versuchsergebnis mit 17,6%iger Wahrscheinlichkeit im Verwerfungsbereich von H_o . Dies ist auch der Fehler 1.Art (α-Fehler) .
$P_{0,2}(X \geq 5) = 74{,}5\%$ Also liegt das Versuchsergebnis mit 74,5%iger Wahrscheinlichkeit im Annahmebereich von H_1 .
$P_{0,2}(X \leq 4) = 25{,}5\%$ Also liegt das Versuchsergebnis mit 25,5%iger Wahrscheinlichkeit im Verwerfungsbereich von H_1 . Dies ist auch der Fehler 2.Art (β-Fehler) .

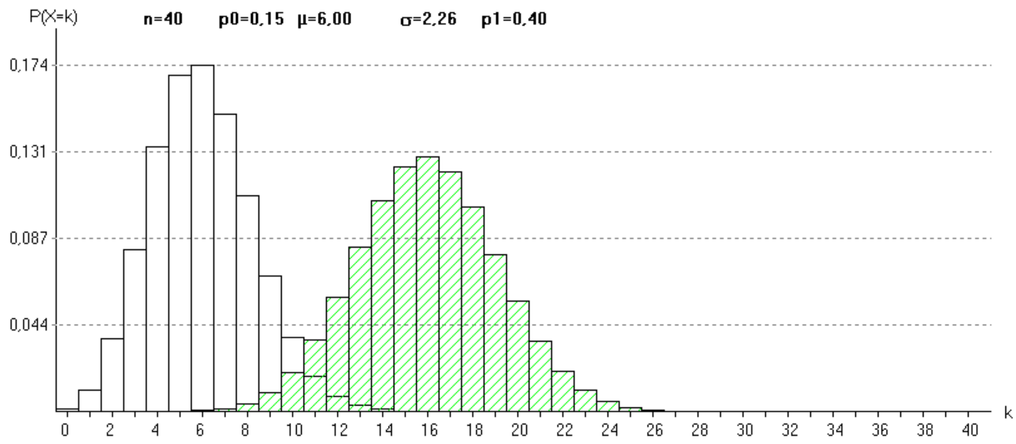
Man sieht, dass die Irrtumswahrscheinlichkeiten α und β hier ziemlich groß sind. Dies liegt daran, dass die beiden Alternativen p=0,1 und p=0,2 so dicht beieinander liegen.

Weitere Beispiele:

Beispiel 1 (Sektkellerei):

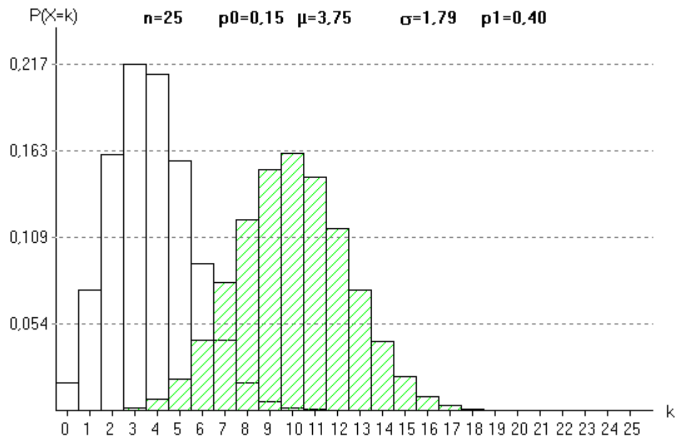
Die einer Sektkellerei gelieferten Flaschen enthalten aufgrund von Produktionsschwankungen entweder zu 15% (1.Qualität) oder zu 40% (2.Qualität) leichte Farbveränderungen.
Der Kellermeister möchte die Qualität einer Lieferung erkennen.

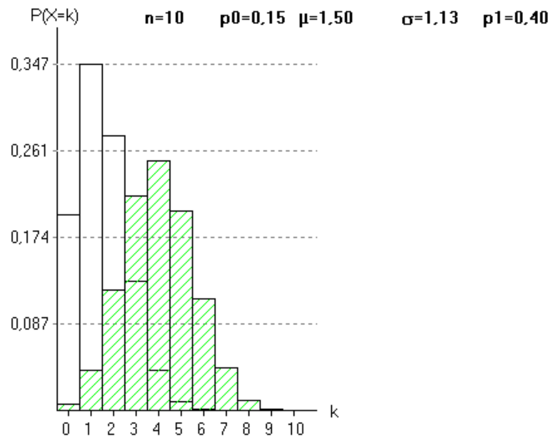
Er entnimmt einer Lieferung 40 Flaschen und prüft sie. Die Anzahl der Flaschen mit Farbveränderung sei k. Wie lautet die Entscheidungsregel ? Mit welchen Wahrscheinlichkeiten kann der Kellermeister sich irren ?



Regel: k≤10 (1.Qualität) k≥11 (2.Qualität)
Irrtumswahrscheinlichkeiten: α-Fehler = 3% β-Fehler = 3,5%

Der Kellermeister könnte aber auch nur 25 (10) Flaschen prüfen . Wie sehen dann die entsprechenden Ergebnisse aus ?





k≤6 (1.Qualität) k≥7 (2.Qualität)

α-Fehler = 6,9% β-Fehler = 7,4%

k≤2 (1.Qualität) k≥3 (2.Qualität)

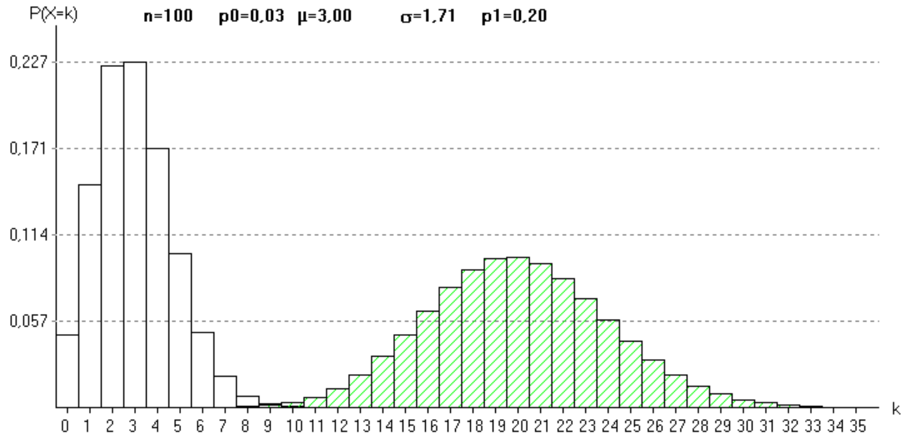
α-Fehler = 18% β-Fehler = 16,7%

<u>Beispiel 2 (Abitur 2006, jedoch abgewandelt):</u>

Eine Firma produziert Roboter, wobei zu Beginn 20% fehlerhaft sind. Nach Umstellung der Produktion soll die Fehlerquote auf 3% reduziert werden. Ein Kunde überprüft zur Kontrolle 100 Roboter und entwirft einen Alternativtest (auf der Grundlage des **Vergleichs der beiden Verteilungen**) um entscheiden zu können, ob er eine Lieferung beanstanden soll oder nicht. Er entdeckt bei der Prüfung 11 defekte Roboter .

<u>Lösungsweg:</u>
Wir nehmen an, dass der Kunde nachweisen will, dass die Fehlerquote immer noch bei 20% liegt. Die **<u>gegenteilige Behauptung ist dann die Nullhypothese</u>** H_o, die es zu widerlegen gilt.
Somit gilt: n=100 H_o: p=0,03 H_1: p=0,2 Stichprobenergebnis: 11

Wir betrachten zunächst die beiden Verteilungen, um eine Grenze zwischen beiden zu bestimmen:



<u>Folgerungen:</u>

1) Der Verwerfungsbereich bzgl. der Nullhypothese H_o: p=0,03 ist (in der Grafik leider schlecht zu erkennen, daher besser die Tabelle mit P0(X=k) und P1(X=k) konsultieren !)
V = {10; … ;100} . Annahmebereich also A = {0; … ;9}

Weil das Stichprobenergebnis 11 in V liegt, wird die **<u>Nullhypothese abgelehnt</u>**, d.h. der Kunde geht davon aus, dass die Fehlerquote sich trotz Produktionsumstellung nicht verbessert hat !

2) Der α-Fehler (das Stichprobenergebnis liegt im Verwerfungsbereich von H_o, obwohl H_o richtig ist) beträgt laut Tabelle: $P_{0,03}(X>9) = 1 - \text{binomcdf}(100,.03,9) \approx 0,09\%$

3) Der β-Fehler (das Stichprobenergebnis liegt im Annahmebereich von H_o, obwohl H_o falsch und somit H_1 richtig ist) beträgt laut Tabelle : $P_{0,2}(X\leq 9) = \text{binomcdf}(100,.2,9) \approx 0,2\%$

<u>Ergänzende Aufgabe zu Beispiel 2:</u>
Obige Betrachtung mit der Festlegung der Nullhypothese berücksichtigt die Sicht des Kunden . Wie würde es aus der Sicht des Herstellers aussehen ??

<u>Lösungsweg:</u>
H_o: p=0,2 H_1: p=0,03 (vertauschte Hypothesen, da der Hersteller p=0,03 beweisen will !).
Ablehnungsbereich für H_o: p=0,2: V = {0; … ;9}. 11 liegt aber in A, also würde auch der Hersteller die Hypothese p=0,2 nicht ablehnen können und somit p=0,03 ablehnen (müssen) !! Er kann also die Verbesserung seiner Produktion nicht nachweisen !

Bei diesen vertauschten Hypothesen (bzgl. der vorherigen Betrachtung) vertauschen sich auch die α- und β-Fehler . $\alpha = 0,2\%$ und $\beta = 0,09\%$.

Beispiel 3 (ähnlich wie Beispiel 2; vgl. Abitur NS 2006; aber mit 5%-Signifikanzniveau):

Eine Firma produziert Roboter, wobei zu Beginn 20% fehlerhaft sind. Nach Umstellung der Produktion soll die Fehlerquote auf 3% reduziert werden.
Ein Kunde überprüft zur Kontrolle 100 Roboter und entwirft ein Testverfahren auf dem Signifikanzniveau von 5% um entscheiden zu können, ob er eine Lieferung beanstanden soll oder nicht. Er entdeckt bei der Prüfung 11 defekte Roboter .

Lösungsweg:
Wir nehmen an, dass der Kunde nachweisen will, dass die Fehlerquote immer noch bei 20% liegt. Die gegenteilige Behauptung ist dann die Nullhypothese H_o, die es zu widerlegen gilt.
Somit gilt: n=100 H_o: p=0,03 H_1: p=0,2 $\alpha \le 5\%$ Stichprobenergebnis: 11

Wir betrachten zunächst die **kumulierten Verteilungen** der Hypothesen (insbesondere P0(X>=k):

k	P0(X=k)	P1(X=k)	P0(X<=k)	P1(X<=k)	P0(X>=k)
0	4,76%	0,00%	4,76%	0,00%	100,00%
1	14,71%	0,00%	19,46%	0,00%	95,24%
2	22,52%	0,00%	41,98%	0,00%	80,54%
3	22,75%	0,00%	64,72%	0,00%	58,02%
4	17,06%	0,00%	81,79%	0,00%	35,28%
5	10,13%	0,00%	91,92%	0,00%	18,21%
6	4,96%	0,01%	96,88%	0,01%	8,08%
7	2,06%	0,02%	98,94%	0,03%	3,12%
8	0,74%	0,06%	99,68%	0,09%	1,06%
9	0,23%	0,15%	99,91%	0,23%	0,32%
10	0,07%	0,34%	99,98%	0,57%	0,09%
11	0,02%	0,69%	100,00%	1,26%	0,02%

Folgerungen:

1) Der Verwerfungsbereich auf dem 5%-Niveau bzgl. der Nullhypothese p=0,03 ist
V = {7; … ;100} . (rechtsseitig !)
Annahmebereich A = {0; … ;6}

2) Weil das Stichprobenergebnis 11 im Bereich V liegt, wird die **Nullhypothese abgelehnt**, d.h. der Kunde geht davon aus, dass die Fehlerquote sich trotz Produktionsumstellung nicht verbessert hat !

3) Der α-Fehler (das Stichprobenergebnis liegt im Verwerfungsbereich von H_o, obwohl H_o richtig ist) beträgt: $P_{0,03}(X>6) \approx 1 - 0,97 = 3\%$

4) Der β-Fehler (das Stichprobenergebnis liegt im Annahmebereich von H_o, obwohl H_o falsch und somit H_1 richtig ist) beträgt laut Tabelle :
$P_{0,2}(X \le 6) = 7,8 \cdot 10^{-5} \approx 0,008\%$

Ergänzende Aufgabe zu Beispiel 3:
Obige Betrachtung mit der Festlegung der Nullhypothese berücksichtigt die Sicht des Kunden .
Wie würde es aus der Sicht des Herstellers aussehen ??

Lösungsweg:

H_o: p=0,2 H_1: p=0,03 (vertauschte Hypothesen, da der Hersteller p=0,03 beweisen will und somit das Gegenteil p=0,2 annimmt !).
Verwerfungsbereich für p=0,2 und α=0,05: V = {0; … ;13}. Also ist A = {14; … ;100} .
11 liegt in V, also würde der Hersteller hier im Gegensatz zu der Kundensicht die Hypothese p=0,2 ablehnen und somit p=0,03 (Produktionsverbesserung) annehmen !!

Überblick **Alternativtest** bei Binomialverteilungen

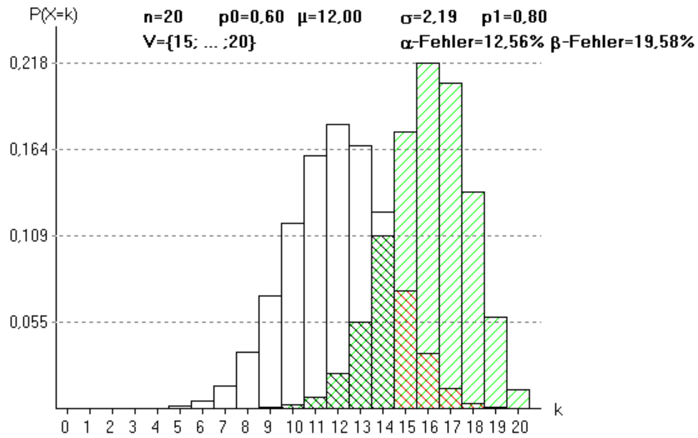
Beim Alternativtest will man mit einer Stichprobe vom Umfang n eine Hypothese H_o (z.B. p_o=70%) widerlegen . Man geht aus von einer Binomialverteilung $B_{n;po}$ (k=0;1; … ;n) .
Das Signifikanzniveau (Verwerfungsbereich V der Nullhypothese) wird entweder mit α (meist $\alpha \leq 5\%$) oder aber mit einer Schranke k (V={0; … ;k} bzw. V= {k; … ;n}) festgelegt.
V ist ein Bereich (für die k-Werte) in der Binomialverteilung $B_{n;po}$ (je nach Lage der Alternativwahrscheinlichkeit p_1 liegt V entweder links oder rechts bezüglich H_o) .

<u>Beispiel:</u> n=20 H_o: p=0,6 H_1: p=0,8
α sei hier nicht vorgegeben, so dass V durch Vergleich der Verteilungen festgelegt wird.

k	P0(X=k)	P1(X=k)	P0(X<=k)	P1(X<=k)	P0(X>=k)
0	0,00%	0,00%	0,00%	0,00%	100,00%
1	0,00%	0,00%	0,00%	0,00%	100,00%
2	0,00%	0,00%	0,00%	0,00%	100,00%
3	0,00%	0,00%	0,00%	0,00%	100,00%
4	0,03%	0,00%	0,03%	0,00%	100,00%
5	0,13%	0,00%	0,16%	0,00%	99,97%
6	0,49%	0,00%	0,65%	0,00%	99,84%
7	1,46%	0,00%	2,10%	0,00%	99,35%
8	3,55%	0,01%	5,65%	0,01%	97,90%
9	7,10%	0,05%	12,75%	0,06%	94,35%
10	11,71%	0,20%	24,47%	0,26%	87,25%
11	15,97%	0,74%	40,44%	1,00%	75,53%
12	17,97%	2,22%	58,41%	3,21%	59,56%
13	16,59%	5,45%	75,00%	8,67%	41,59%
14	12,44%	10,91%	87,44%	19,58%	25,00%
15	7,46%	17,46%	94,90%	37,04%	12,56%
16	3,50%	21,82%	98,40%	58,86%	5,10%
17	1,23%	20,54%	99,64%	79,39%	1,60%
18	0,31%	13,69%	99,95%	93,08%	0,36%
19	0,05%	5,76%	100,00%	98,85%	0,05%
20	0,00%	1,15%	100,00%	100,00%	0,00%

Die Tabelle zeigt, dass P1(X=k) von k=15 ab größere Wahrscheinlichkeiten besitzt als P0(X=k).

Daher verwerfen wir H_0, wenn das Stichprobenergebnis größer oder gleich 15 ist.



<u>Entscheidungsregel:</u> Lehne H_o ab, wenn das Stichprobenergebnis in V = {15 ; … ; 20} liegt.
(15 ergibt sich daraus, dass ab k=15 H_1 größere Wahrscheinlichkeiten besitzt als H_0)

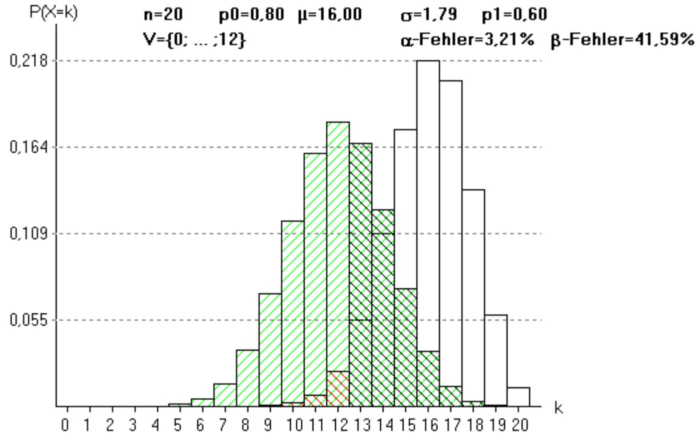
Fehler 1.Art (α-Fehler): Eine richtige Hypothese wird abgelehnt. D.h.: das Stichprobenergebnis liegt im Verwerfungsbereich V (rot schraffiert), obwohl H_o richtig ist .
In obigem Beispiel gilt: α-Fehler = $P_{0,6}(X \geq 15) = 1 - P_{0,6}(X \leq 14) = 12,6\%$

Fehler 2.Art (β-Fehler): Eine falsche Hypothese wird beibehalten . D.h.: das Stichprobenergebnis liegt im Annahmebereich A (schwarz schraffiert), obwohl H_o falsch ist .
Um den β-Fehler berechnen zu können muss die richtige Wahrscheinlichkeit p_1 bekannt sein.
In obigem Beispiel gilt für p_1=0,8: β-Fehler = $P_{0,8}(X \leq 14) = 19,6\%$

Alternativtest bei prozentual festgelegtem V

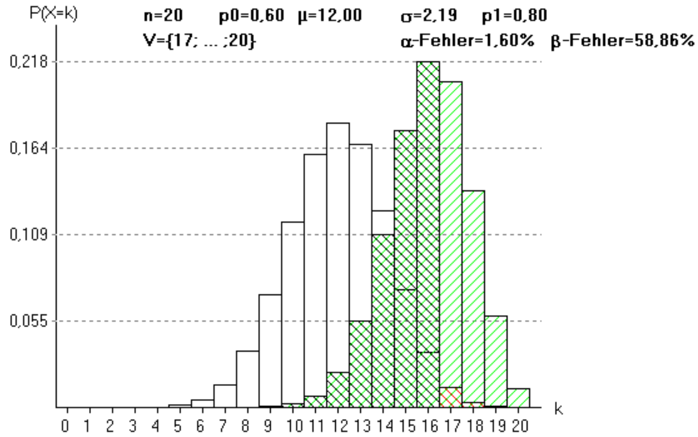
Es ist in der Regel so, dass der Verwerfungsbereich V nicht durch Vergleich der Verteilungen, sondern durch eine bestimmte %-Zahl festgelegt wird, z.B. $\alpha \leq 5\%$.
Man spricht dann von einem Signifikanzniveau von 5% .
Betrachtet werden muss hier nun die <u>kumulative</u> Wahrscheinlichkeitstabelle für p0.
Liegt V rechtsseitig, so lässt es sich schwerer bestimmen als wenn es linksseitig liegt !

Beispiel1: (V linksseitig) n=20 H_o: p=0,8 H_1: p=0,6 $\alpha \leq 5\%$



Der Verwerfungsbereich ergibt sich hier ersichtlich zu V={0; … ;12}, was man auch aus der Tabelle für P0(X≤k) ablesen kann .

Beispiel2: (V rechtsseitig) n=20 H_o: p=0,6 H_1: p=0,8 $\alpha \leq 5\%$



Der Verwerfungsbereich ergibt sich hier ersichtlich zu V={17; … ;20}, was man aber aus der Tabelle für P0(X≤k) nur schwer ablesen kann .

P0(X≤k) muss ≥ 95% sein .

Aufgaben:

1) Bestimme V für n=20 p_o=0,6 p_1=0,8 $\alpha \leq 1\%$ Lösung: V={18; … ;20}

2) Bestimme V für n=20 p_o=0,6 p_1=0,8 $\alpha \leq 10\%$ Lösung: V={16; … ;20}

3) Wie lautet der Verwerfungsbereich V für n=50 p_o=0,4 p_1=0,1 $\alpha \leq 5\%$?
Berechne auch die Fehler 1. und 2. Art .

Lösung: V={0; … ;13} α-Fehler = $P_{0,4}(X \leq 13) = 2{,}8\%$
β-Fehler = $P_{0,1}(X \geq 14) = 1- P_{0,1}(X \leq 13) = 0{,}03\%$

4) Wie lautet der Verwerfungsbereich V für n=50 p_o=0,4 p_1=0,7 $\alpha \leq 5\%$?
Berechne auch die Fehler 1. und 2. Art .

Lösung: V={27; … ;50} α-Fehler = $P_{0,4}(X \geq 27) = 1- P_{0,4}(X \leq 26) = 3{,}14\%$
β-Fehler = $P_{0,7}(X \leq 26) = 0{,}56\%$

Man kann die Zufallsversuche mit dem GTR auch **simulieren**.
So erhält man ein oder mehrere Stichprobenergebnisse .

Simulationsbeispiel (für $n=20$ $p_o=0,6$ $p_1=0,8$):

Der Verwerfungsbereich ist $V=\{15 ; \dots ; 20\}$ (siehe oben)
Der Annahmebereich ist also $A=\{0 ; \dots ; 14\}$

Wir simulieren mehrere Zufallsversuche mit dem Befehl **randBin** des GTR, so dass wir N Stichprobenergebnisse erhalten.

Der Aufruf von z.B. randBin(20,0.6,15) liefert 10 Stichprobenergebnisse:

Stichprobenergebnis k	**Entscheidung**
15	Da $15 \in V$ wird die Hypothese H_o verworfen
9	Da $9 \in A$ wird die Hypothese H_o angenommen
10	H_o angenommen
14	H_o angenommen
14	H_o angenommen
11	H_o angenommen
8	H_o angenommen
10	H_o angenommen
13	H_o angenommen
11	H_o angenommen
12	H_o angenommen
11	H_o angenommen
7	H_o angenommen
3	H_o angenommen
18	H_o verworfen

Erweiterung:

Lässt man den GTR nun z.B. 500 Simulationen durchführen, so kann man die relative Häufigkeit für Stichprobenergebnisse ≥ 15 als Schätzwert für den α-Fehler ansehen:

randBin(20,.6,500) STO L_1	erstellt die Liste der Stichprobenergebnisse
sum(L1≥ 15)	liefert den Wert H=78 für die abs.H. des α-Fehlers.
Also ist die rel.H. des α-Fehlers	$h = 78/500 = 15,6\%$ (schlechter Schätzwert !)